GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 18, 2003, 07:42:15 ; Search time 21 Seconds
(without alignments)
767.640 Million cell updates/sec

Title: US-09-623-035-2
Perfect score: 2064
Sequence: 1 MTVARPSVPAALPLLGELPR.....HTCFTLTGLLGTLVTMGLLT 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 157478

Minimum DB seq length: 6
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	124	6.0	23	1	US-07-811-048-8	Sequence 8, Appli
2	96	4.7	27	6	5256642-18	Patent No. 5256642
3	96	4.7	27	6	5472939-18	Patent No. 5472939
4	50.5	2.4	27	2	US-08-541-759B-5	Sequence 5, Appli
5	48	2.3	26	1	US-08-450-360-12	Sequence 12, Appl
6	47.5	2.3	27	2	US-08-541-759B-4	Sequence 4, Appli
7	47.5	2.3	29	4	US-09-043-731-23	Sequence 23, Appl
8	47	2.3	8	1	US-08-210-266A-4	Sequence 4, Appli
9	47	2.3	8	1	US-08-688-675-4	Sequence 4, Appli
10	47	2.3	8	3	US-08-477-860C-4	Sequence 4, Appli
11	47	2.3	27	2	US-08-169-948B-22	Sequence 22, Appl
12	47	2.3	27	2	US-08-448-873-22	Sequence 22, Appl
13	47	2.3	27	3	US-08-382-452D-22	Sequence 22, Appl
14	47	2.3	27	4	US-08-507-362A-10	Sequence 10, Appl
15	46.5	2.3	26	2	US-08-288-059-2	Sequence 2, Appli
16	46	2.2	27	2	US-08-541-759B-3	Sequence 3, Appli
17	45	2.2	10	6	5378464-24	Patent No. 5378464
18	45	2.2	22	6	5378464-8	Patent No. 5378464
19	45	2.2	24	3	US-08-701-382-8	Sequence 8, Appli
20	45	2.2	24	4	US-08-788-820-8	Sequence 8, Appli
21	44.5	2.2	25	1	US-07-966-187-4	Sequence 4, Appli
22	44.5	2.2	25	1	US-08-371-121-3	Sequence 3, Appli
23	44.5	2.2	25	3	US-07-927-391-5	Sequence 5, Appli
24	44.5	2.2	25	5	PCT-US94-04361-32	Sequence 32, Appl
25	44	2.1	22	4	US-09-461-325-488	Sequence 488, App
26	44	2.1	25	5	PCT-US94-04361-33	Sequence 33, Appl
27	43.5	2.1	22	6	5256642-15	Patent No. 5256642
28	43.5	2.1	22	6	5472939-15	Patent No. 5472939
29	43.5	2.1	23	2	US-08-833-807-2	Sequence 2, Appli
30	43.5	2.1	23	3	US-09-223-043-2	Sequence 2, Appli
31	43.5	2.1	23	4	US-09-593-870A-2	Sequence 2, Appli
32	43	2.1	14	6	5256642-19	Patent No. 5256642
33	43	2.1	14	6	5472939-19	Patent No. 5472939
34	43	2.1	19	2	US-08-648-298-16	Sequence 16, Appl
35	43	2.1	25	4	US-09-439-313-520	Sequence 520, App
36	43	2.1	27	3	US-09-071-710-39	Sequence 39, Appl
37	43	2.1	27	3	US-09-525-397-39	Sequence 39, Appl
38	43	2.1	27	4	US-09-439-313-566	Sequence 566, App
39	43	2.1	29	4	US-09-439-313-546	Sequence 546, App
40	42	2.0	7	1	US-08-210-266A-5	Sequence 5, Appli
41	42	2.0	7	1	US-08-688-675-5	Sequence 5, Appli
42	42	2.0	7	3	US-08-477-860C-5	Sequence 5, Appli
43	41.5	2.0	16	1	US-08-151-219-1	Sequence 1, Appli
44	41.5	2.0	16	5	PCT-US94-13205-1	Sequence 1, Appli
45	41.5	2.0	25	5	PCT-US94-04361-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-07-811-048-8
; Sequence 8, Application US/07811048
; Patent No. 5264357 ✓ 11-1993
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W.
; TITLE OF INVENTION: Nucleic Acid and Methods for the Synthesis of No. 5264357el Fu
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/811,048
; FILING DATE: 19911219
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/738171
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/859107
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US86/01177
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/083757
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Robert H.
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: 330p1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1489
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-811-048-8

Query Match 6.0%; Score 124; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         35 DCGLPPDVPNAQPALEGRTSFP 56
              ||||||||||||||||||||||
Db          1 DCGLPPDVPNAQPALEGRTSFP 22
```

1, 11, 13, 34
US 5,264,357

```
RESULT 2
5256642-18
;Patent No. 5256642
;     APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
;     TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
;RECEPTOR 1 (CR1) AND A THROMBOLTIC AGENT, AND THE METHODS OF
;USE THEREOF
;     NUMBER OF SEQUENCES: 30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/588,128
;       FILING DATE: 24-SEP-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 412,745
;       FILING DATE: 26-SEP-1989
;       APPLICATION NUMBER: 332,865
;       FILING DATE: 03-APR-1989
;       APPLICATION NUMBER: 176,532
;       FILING DATE: 01-APR-1988
;SEQ ID NO:18:
;       LENGTH: 27
5256642-18

  Query Match             4.7%;  Score 96;  DB 6;  Length 27;
  Best Local Similarity   61.5%;  Pred. No. 0.017;
  Matches   16;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        257 FTMIGEHSIYCTVNNDEGEWSGPPPE 282
              | ::|| ||||| |:| | |||| |:
Db          2 FELVGEPSIYCTSNDDVGIWSGPAPQ 27


RESULT 3
5472939-18
;Patent No. 5472939
;     APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
;     TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
;MEDIATED DISORDERS
;     NUMBER OF SEQUENCES: 30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/138,825
;       FILING DATE: 19-OCT-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 588,128
;       FILING DATE: 24-SEP-1990
;       APPLICATION NUMBER: 412,745
;       FILING DATE: 26-SEP-1989
;       APPLICATION NUMBER: 332,865
;       FILING DATE: 03-APR-1989
;       APPLICATION NUMBER: 176,532
;       FILING DATE: 01-APR-1988
;SEQ ID NO:18:
;       LENGTH: 27
5472939-18

  Query Match             4.7%;  Score 96;  DB 6;  Length 27;
  Best Local Similarity   61.5%;  Pred. No. 0.017;
  Matches   16;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        257 FTMIGEHSIYCTVNNDEGEWSGPPPE 282
              | ::|| ||||| |:| | |||| |:
Db          2 FELVGEPSIYCTSNDDVGIWSGPAPQ 27


RESULT 4
US-08-541-759B-5
; Sequence 5, Application US/08541759B
; Patent No. 5861160
; GENERAL INFORMATION:
;   APPLICANT:  Quick, Douglas P.
;   APPLICANT:  Welter, Mark W.
;   APPLICANT:  Welter, Joseph
;   APPLICANT:  Welter, Lisa M.
;   TITLE OF INVENTION:  ISOSPORA SUIS VACCINE
;   NUMBER OF SEQUENCES:  11
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Akzo No. 5861160el Patent Dept.
;     STREET:  1300 Piccard Drive, Suite 206
;     CITY:  Rockville
;     STATE:  Maryland
;     COUNTRY:  USA
;     ZIP:  20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/541,759B
;     FILING DATE:  10-OCT-1995
;     CLASSIFICATION:  424
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Klesner, Sharon N.
;     REGISTRATION NUMBER:  36,335
;     REFERENCE/DOCKET NUMBER:  Quick1a
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  301-948-7400
;     TELEFAX:  301-948-9751
; INFORMATION FOR SEQ ID NO:  5:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  27 amino acids
;     TYPE:  amino acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  peptide
US-08-541-759B-5

  Query Match             2.4%;  Score 50.5;  DB 2;  Length 27;
  Best Local Similarity   41.4%;  Pred. No. 3.3e+02;
  Matches   12;  Conservative    4;  Mismatches    8;  Indels    5;  Gaps    1;

Qy        292 VPPTVQKPTTVNVPTTEVSPTSQKTTTKT 320
              |||| : |     || ||:| ::  | |
Db          1 VPPTEEVP-----PTEEVTPPTEGETPPT 24


RESULT 5
US-08-450-360-12
; Sequence 12, Application US/08450360
; Patent No. 5656457
; GENERAL INFORMATION:
;   APPLICANT:  Parkes, Deborah Lynn
;   APPLICANT:  Coates, Stephen Ralph
;   TITLE OF INVENTION:  Herpes Simplex Virus Type 2-Glycoprotein G
;   TITLE OF INVENTION:  Proteins and Polypeptides
;   NUMBER OF SEQUENCES:  12
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Leona L. Lauder
;     STREET:  6 Mariposa Court
;     CITY:  Tiburon
;     STATE:  California
;     COUNTRY:  USA
;     ZIP:  94920
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 18, 2003, 07:42:11 ; Search time 49 Seconds
(without alignments)
1234.180 Million cell updates/sec

Title: US-09-623-035-2
Perfect score: 2064
Sequence: 1 MTVARPSVPAALPLLGELPR.....HTCFTLTGLLGTLVTMGLLT 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 436344

Minimum DB seq length: 6
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	147	7.1	29	15	AAR50086	Decay accelerating
2	89.5	4.3	28	23	AAM50805	Human complement r
3	64.5	3.1	23	18	AAW31838	Peptide 4 from the
4	61	3.0	10	22	AAG97023	Human complementar
5	58	2.8	10	22	AAG97081	Human complementar
6	58	2.8	10	22	AAG97083	Human complementar
7	57	2.8	10	22	AAG97043	Human complementar
8	55	2.7	10	22	AAG97037	Human complementar
9	55	2.7	10	22	AAG97045	Human complementar
10	54	2.6	10	22	AAG97089	Human complementar
11	52.5	2.5	19	21	AAY65634	Oestrogen receptor
12	52.5	2.5	19	23	AAU86371	Oestrogen receptor
13	52	2.5	10	22	AAG96997	Human complementar
14	52	2.5	10	22	AAG97001	Human complementar
15	52	2.5	10	22	AAG97015	Human complementar
16	52	2.5	10	22	AAG97059	Human complementar
17	52	2.5	10	22	AAG97091	Human complementar
18	51.5	2.5	27	24	ABU07673	Human haptoglobin
19	51	2.5	10	22	AAG96999	Human complementar
20	51	2.5	10	22	AAG97003	Human complementar
21	51	2.5	10	22	AAG97005	Human complementar
22	51	2.5	10	22	AAG97007	Human complementar
23	51	2.5	10	22	AAG97009	Human complementar
24	51	2.5	10	22	AAG97011	Human complementar
25	51	2.5	10	22	AAG97013	Human complementar
26	51	2.5	10	22	AAG97017	Human complementar
27	51	2.5	10	22	AAG97031	Human complementar
28	51	2.5	10	22	AAG97033	Human complementar
29	51	2.5	10	22	AAG97041	Human complementar
30	51	2.5	10	22	AAG97049	Human complementar
31	51	2.5	10	22	AAG97051	Human complementar
32	51	2.5	10	22	AAG97053	Human complementar
33	51	2.5	10	22	AAG97055	Human complementar
34	51	2.5	10	22	AAG97069	Human complementar
35	51	2.5	10	22	AAG97071	Human complementar
36	51	2.5	10	22	AAG97073	Human complementar
37	51	2.5	10	22	AAG97075	Human complementar
38	51	2.5	10	22	AAG97077	Human complementar
39	51	2.5	10	22	AAG97079	Human complementar
40	51	2.5	10	22	AAG97085	Human complementar
41	51	2.5	10	22	AAG97093	Human complementar
42	51	2.5	10	22	AAG97095	Human complementar
43	51	2.5	10	22	AAG97097	Human complementar
44	51	2.5	10	22	AAG97105	Human complementar
45	51	2.5	10	22	AAG97107	Human complementar

ALIGNMENTS

RESULT 1
AAR50086
ID AAR50086 standard; Protein; 29 AA.
XX
AC AAR50086;
XX
DT 25-MAR-2003 (updated)
DT 27-OCT-1994 (first entry)
XX
DE Decay accelerating factor derived attachment signal.
XX
KW Fusion protein; xenograft; transplantation; membrane; transplant;
KW membrane cofactor protein; MCP; decay accelerating factor; DAF; GPI;
KW glycophosphatidyl inositol; anchor; intermembrane transfer.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 13..29
FT /note= "C-terminal hydrophobic domain."
XX
PN WO9406903-A1.
XX
PD 31-MAR-1994.
XX
PF 22-SEP-1993; 93WO-US08889.
XX
PR 22-SEP-1992; 92US-0948521.
XX
PA (DNXB-) DNX BIOTHERAPEUTICS INC.
XX
PI Byrne GW, Kooyman DL, Logan JS;

XX
DR WPI; 1994-118454/14.
XX
PT Delivery of proteins, using mobile cell, by intermembrane
PT transfer - for pre-accommodation of xenogeneic organ transplants
XX
PS Disclosure; Page 43; 104pp; English.
XX
CC A protein of interest may be delivered to the tissue of a target
CC animal by providing mobile cells which bear the protein on their
CC membranes, linked to a glycophosphatidyl inositol (GPI) anchor. The
CC cells circulate in the target animal, coming into contact with the
CC tissue and transfer the protein to the tissue by intermembrane
CC transfer. Intermembrane transfer by this method occurs at a greater
CC rate than naturally in the animal. This sequence is an attachment
CC signal derived from decay accelerating factor (DAF). DAF is
CC naturally GPI linked. Preferably, the protein of interest is
CC expressed as a fusion product, the fusion protein comprising this
CC sequence and the protein (or fragment) of interest. The method is
CC especially useful for modifying pig organs for xenotransplantation
CC into humans. See AAQ58894.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 29 AA;

Query Match 7.1%; Score 147; DB 15; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 353 SGTTRLLSGHTCFTLTGLLGTLVTMGLLT 381
||||||||||||||||||||||||||||||
Db 1 SGTTRLLSGHTCFTLTGLLGTLVTMGLLT 29

1, 11, 13, 34

RESULT 2
AAM50805
ID AAM50805 standard; Peptide; 28 AA.
XX
AC AAM50805;
XX
DT 01-MAY-2002 (first entry)
XX
DE Human complement receptor type 1 tryptic peptide 54d.
XX
KW Complement receptor type 1; CR1; C3b/C4b receptor; human;
KW inflammation; myocardial infarction; reperfusion injury;
KW antiinflammatory; therapy; diagnosis; receptor; purification.
XX
OS Homo sapiens.
XX
PN US6316604-B1.
XX
PD 13-NOV-2001.
XX
PF 05-JUN-1995; 95US-0463959.
XX
PR 03-APR-1989; 89US-0332865.
PR 06-DEC-1994; 94US-0350238.
PR 24-FEB-1993; 93US-0026134.
PR 01-APR-1988; 88US-0176532.
XX
PA (AVAN-) AVANT IMMUNOTHERAPEUTICS INC.
XX
PI Fearon DT, Klickstein LB, Wong WW, Carson GR, Concino MF, Ip SH;
PI Makrides SC, Marsh HC;
XX
DR WPI; 2002-121028/16.
XX
Purifying a recombinant soluble complement receptor type 1 molecule or
fragment, useful for diagnosing or treating disorders associated
plement activity, using cation exchange chromatography provides
PS Example; Column 25; 44pp; English.
XX
CC The present sequence is that of tryptic peptide 54d of human
CC erythrocyte C3b/C4b receptor CR1 (complement receptor type 1).
CC The peptide corresponds to amino acids 204-231 and 654-681 of a
CC full-length CR1 sequence (see AAM50797) predicted from isolated cDNA
CC clones. The invention provides recombinant, soluble CR1 polypeptides
CC that lack the transmembrane domain of the native protein, and methods
CC of purifying them. CR1 polypeptides and nucleic acids have use in
CC the diagnosis and therapy of disorders involving complement activity
CC and various immune system or inflammatory disorders. Secreted,
CC recombinant CR1 has been produced in large quantities using CHO host
CC cells grown in serum-free medium. The purified protein was used to
CC reduce damage caused by inflammation, and to reduce myocardial
CC infarct size and prevent reperfusion injury in rats.
XX
SQ Sequence 28 AA;

Query Match 4.3%; Score 89.5; DB 23; Length 28;
Best Local Similarity 59.3%; Pred. No. 3.9;
Matches 16; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 257 FTMIGEHSIYCTVNNDE-GEWSGPPPE 282
| ::|| ||||| |:|: | |||| |:
Db 2 FELVGEPSIYCTSNDDQVGIWSGPAPQ 28

RESULT 3
AAW31838
ID AAW31838 standard; peptide; 23 AA.
XX
AC AAW31838;
XX
DT 26-MAR-1998 (first entry)
XX
DE Peptide 4 from the short consensus repeat 3 of complement receptor 1.
XX
KW Short consensus repeat 3; long homologous repeat A; LHR-A;
KW complement receptor 1; CR1; complement inhibition;
KW anti-haemolytic activity; inflammation; thrombotic condition;
KW inappropriate complement activation; ARDS; Alzheimer's disease.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9731944-A1.
XX
PD 04-SEP-1997.
XX
PF 26-FEB-1997; 97WO-EP00994.
XX
PR 02-MAR-1996; 96GB-0004518.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Edge CM, Mossakowska DEI, Smith RAG;
XX
DR WPI; 1997-448630/41.
XX
PT Peptide derived from short consensus repeat 3 of human complement
PT receptor 1 - between amino acids Cys154-Gly186, useful to treat
PT disorder or disease associated with inflammation or inappropriate
PT complement activation
XX
PS Claim 13; Page 25; 33pp; English.
XX
CC Synthetic soluble peptides AAW31835-38 are derived from the short
CC consensus repeat 3 of the long homologous repeat A (LHR-A) of human
CC complement receptor 1 (CR1). The present peptide is located between
CC residues 164 and 186. These peptides have functional complement
CC inhibitory, including anti-haemolytic, activity. Enhanced activity
CC may be achieved by linking the peptides to a core structure
CC (e.g. the MAP peptide) to produce a multimeric or chimeric polypeptide.

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,360
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/129,021
; FILING DATE: 29-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0012.2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; GENERAL INFORMATION:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-450-360-12

Query Match 2.3%; Score 48; DB 1; Length 26;
Best Local Similarity 41.7%; Pred. No. 5.4e+02;
Matches 10; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

Qy 311 PTSQKTTTKTTTPNAQATRSTPVS 334
||| | : ||| | | | :
Db 1 PTSTHATPRPTTPGPQTTPPGPAT 24

RESULT 6
US-08-541-759B-4
; Sequence 4, Application US/08541759B
; Patent No. 5861160
; GENERAL INFORMATION:
; APPLICANT: Quick, Douglas P.
; APPLICANT: Welter, Mark W.
; APPLICANT: Welter, Joseph
; APPLICANT: Welter, Lisa M.
; TITLE OF INVENTION: ISOSPORA SUIS VACCINE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5861160el Patent Dept.
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/541,759B
; FILING DATE: 10-OCT-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Klesner, Sharon N.
; REGISTRATION NUMBER: 36,335
; REFERENCE/DOCKET NUMBER: Quick1a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-948-7400
; TELEFAX: 301-948-9751
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-541-759B-4

Query Match 2.3%; Score 47.5; DB 2; Length 27;
Best Local Similarity 37.9%; Pred. No. 6.3e+02;
Matches 11; Conservative 5; Mismatches 8; Indels 5; Gaps 1;

Qy 292 VPPTVQKPTTVNVPTTEVSPTSQKTTTKT 320
:||| : | || ||:| :: | |
Db 1 LPPTEEVP-----PTEEVTPPTEGETPPT 24

RESULT 7
US-09-043-731-23
; Sequence 23, Application US/09043731A
; Patent No. 6344203
; GENERAL INFORMATION:
; APPLICANT: The Austin Research Institute
; TITLE OF INVENTION: Mimicking Peptides in Cancer Therapy
; FILE REFERENCE: CALA-200
; CURRENT APPLICATION NUMBER: US/09/043,731A
; CURRENT FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: single
; OTHER INFORMATION: stranded linear peptide
US-09-043-731-23

Query Match 2.3%; Score 47.5; DB 4; Length 29;
Best Local Similarity 48.0%; Pred. No. 7e+02;
Matches 12; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

Qy 299 PTTVNVPTTE-VSPTSQKTTTKTTT 322
||| : || |:|| | | :| |
Db 3 PTTTPISTTTMVTPTPTPTGTQTPT 27

RESULT 8
US-08-210-266A-4
; Sequence 4, Application US/08210266A
; Patent No. 5545619 17 08-1996
; GENERAL INFORMATION:
; APPLICANT: Atkinson, John P.
; APPLICANT: Hourcade, Dennis
; APPLICANT: Krych, Malgorzata
; TITLE OF INVENTION: Modified Complement System
; TITLE OF INVENTION: Regulators
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/210,266A
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/695,514
; FILING DATE: 03-MAY-1991
; ATTORNEY/AGENT INFORMATION:

; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)873-8794
; TELEFAX: (404)873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-210-266A-4

Query Match 2.3%; Score 47; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 SDPLPECR 221
||||||||
Db 1 SDPLPECR 8

1, 34

RESULT 9
US-08-688-675-4
; Sequence 4, Application US/08688675
; Patent No. 5719127
; GENERAL INFORMATION:
; APPLICANT: Atkinson, John P.
; APPLICANT: Hourcade, Dennis
; APPLICANT: Krych, Malgorzata
; TITLE OF INVENTION: Modified Complement System Regulators
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,675
; FILING DATE: 30-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/210,266
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/695,514
; FILING DATE: 03-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU101div2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)873-8794
; TELEFAX: (404)873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-688-675-4

2.3%; Score 47; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 SDPLPECR 221
||||||||
Db 1 SDPLPECR 8

RESULT 10
US-08-477-860C-4
; Sequence 4, Application US/08477860C
; Patent No. 6010873
; GENERAL INFORMATION:
; APPLICANT: Atkinson, John P.
; APPLICANT: Hourcade, Dennis
; APPLICANT: Krych, Malgorzata
; TITLE OF INVENTION: Modified Complement System Regulators
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,860C
; FILING DATE: 7-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/210,266
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/695,514
; FILING DATE: 03-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU 101 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)873-8794
; TELEFAX: (404)873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-477-860C-4

Query Match 2.3%; Score 47; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 SDPLPECR 221
||||||||
Db 1 SDPLPECR 8

RESULT 11
US-08-169-948B-22
; Sequence 22, Application US/08169948B
; Patent No. 5861271
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; APPLICANT: Ward, Michael
; APPLICANT: Clarkson, Kathleen

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 18, 2003, 07:42:15 ; Search time 21 Seconds
(without alignments)
767.640 Million cell updates/sec

Title: US-09-623-035-2
Perfect score: 2064
Sequence: 1 MTVARPSVPAALPLLGELPR.....HTCFTLTGLLGTLVTMGLLT 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 157478

Minimum DB seq length: 6
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	124	6.0	23	1	US-07-811-048-8	Sequence 8, Appli
2	96	4.7	27	6	5256642-18	Patent No. 5256642
3	96	4.7	27	6	5472939-18	Patent No. 5472939
4	50.5	2.4	27	2	US-08-541-759B-5	Sequence 5, Appli
5	48	2.3	26	1	US-08-450-360-12	Sequence 12, Appl
6	47.5	2.3	27	2	US-08-541-759B-4	Sequence 4, Appli
7	47.5	2.3	29	4	US-09-043-731-23	Sequence 23, Appl
8	47	2.3	8	1	US-08-210-266A-4	Sequence 4, Appli
9	47	2.3	8	1	US-08-688-675-4	Sequence 4, Appli
10	47	2.3	8	3	US-08-477-860C-4	Sequence 4, Appli
11	47	2.3	27	2	US-08-169-948B-22	Sequence 22, Appl
12	47	2.3	27	2	US-08-448-873-22	Sequence 22, Appl
13	47	2.3	27	3	US-08-382-452D-22	Sequence 22, Appl
14	47	2.3	27	4	US-08-507-362A-10	Sequence 10, Appl
15	46.5	2.3	26	2	US-08-288-059-2	Sequence 2, Appli
16	46	2.2	27	2	US-08-541-759B-3	Sequence 3, Appli
17	45	2.2	10	6	5378464-24	Patent No. 5378464
18	45	2.2	22	6	5378464-8	Patent No. 5378464
19	45	2.2	24	3	US-08-701-382-8	Sequence 8, Appli
20	45	2.2	24	4	US-08-788-820-8	Sequence 8, Appli
21	44.5	2.2	25	1	US-07-966-187-4	Sequence 4, Appli
22	44.5	2.2	25	1	US-08-371-121-3	Sequence 3, Appli
23	44.5	2.2	25	3	US-07-927-391-5	Sequence 5, Appli
24	44.5	2.2	25	5	PCT-US94-04361-32	Sequence 32, Appl
25	44	2.1	22	4	US-09-461-325-488	Sequence 488, App
26	44	2.1	25	5	PCT-US94-04361-33	Sequence 33, Appl
27	43.5	2.1	22	6	5256642-15	Patent No. 5256642
28	43.5	2.1	22	6	5472939-15	Patent No. 5472939
29	43.5	2.1	23	2	US-08-833-807-2	Sequence 2, Appli
30	43.5	2.1	23	3	US-09-223-043-2	Sequence 2, Appli
31	43.5	2.1	23	4	US-09-593-870A-2	Sequence 2, Appli
32	43	2.1	14	6	5256642-19	Patent No. 5256642
33	43	2.1	14	6	5472939-19	Patent No. 5472939
34	43	2.1	19	2	US-08-648-298-16	Sequence 16, Appl
35	43	2.1	25	4	US-09-439-313-520	Sequence 520, App
36	43	2.1	27	3	US-09-071-710-39	Sequence 39, Appl
37	43	2.1	27	3	US-09-525-397-39	Sequence 39, Appl
38	43	2.1	27	4	US-09-439-313-566	Sequence 566, App
39	43	2.1	29	4	US-09-439-313-546	Sequence 546, App
40	42	2.0	7	1	US-08-210-266A-5	Sequence 5, Appli
41	42	2.0	7	1	US-08-688-675-5	Sequence 5, Appli
42	42	2.0	7	3	US-08-477-860C-5	Sequence 5, Appli
43	41.5	2.0	16	1	US-08-151-219-1	Sequence 1, Appli
44	41.5	2.0	16	5	PCT-US94-13205-1	Sequence 1, Appli
45	41.5	2.0	25	5	PCT-US94-04361-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-07-811-048-8
; Sequence 8, Application US/07811048
; Patent No. 5264357
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W.
; TITLE OF INVENTION: Nucleic Acid and Methods for the Synthesis of No. 5264357el Fus
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech).
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/811,048
; FILING DATE: 19911219
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/738171
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/859107
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US86/01177
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/083757
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Robert H.
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: 330p1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1489
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-811-048-8

Query Match 6.0%; Score 124; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1, 11, 13, 34

```
Qy         35 DCGLPPDVPNAQPALEGRTSFP 56
              ||||||||||||||||||||||
Db          1 DCGLPPDVPNAQPALEGRTSFP 22
```

RESULT 2
5256642-18
;Patent No. 5256642
; APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
;RECEPTOR 1 (CR1) AND A THROMBOLTIC AGENT, AND THE METHODS OF
;USE THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/588,128
;   FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 412,745
;   FILING DATE: 26-SEP-1989
;   APPLICATION NUMBER: 332,865
;   FILING DATE: 03-APR-1989
;   APPLICATION NUMBER: 176,532
;   FILING DATE: 01-APR-1988
;SEQ ID NO:18:
; LENGTH: 27
5256642-18

  Query Match 4.7%; Score 96; DB 6; Length 27;
  Best Local Similarity 61.5%; Pred. No. 0.017;
  Matches 16; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Qy        257 FTMIGEHSIYCTVNNDEGEWSGPPPE 282
              | ::|| ||||| |:| | |||| |:
Db          2 FELVGEPSIYCTSNDDVGIWSGPAPQ 27
```

RESULT 3
5472939-18
;Patent No. 5472939
; APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
;MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/138,825
;   FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 588,128
;   FILING DATE: 24-SEP-1990
;   APPLICATION NUMBER: 412,745
;   FILING DATE: 26-SEP-1989
;   APPLICATION NUMBER: 332,865
;   FILING DATE: 03-APR-1989
;   APPLICATION NUMBER: 176,532
;   FILING DATE: 01-APR-1988
;SEQ ID NO:18:
; LENGTH: 27
5472939-18

  Query Match 4.7%; Score 96; DB 6; Length 27;
  Best Local Similarity 61.5%; Pred. No. 0.017;
  Matches 16; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Qy        257 FTMIGEHSIYCTVNNDEGEWSGPPPE 282
              | ::|| ||||| |:| | |||| |:
Db          2 FELVGEPSIYCTSNDDVGIWSGPAPQ 27
```

RESULT 4
US-08-541-759B-5
; Sequence 5, Application US/08541759B
; Patent No. 5861160
; GENERAL INFORMATION:
;   APPLICANT: Quick, Douglas P.
;   APPLICANT: Welter, Mark W.
;   APPLICANT: Welter, Joseph
;   APPLICANT: Welter, Lisa M.
;   TITLE OF INVENTION: ISOSPORA SUIS VACCINE
;   NUMBER OF SEQUENCES: 11
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Akzo No. 5861160el Patent Dept.
;   STREET: 1300 Piccard Drive, Suite 206
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/541,759B
;   FILING DATE: 10-OCT-1995
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Klesner, Sharon N.
;   REGISTRATION NUMBER: 36,335
;   REFERENCE/DOCKET NUMBER: Quick1a
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 301-948-7400
;   TELEFAX: 301-948-9751
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 27 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-541-759B-5

  Query Match 2.4%; Score 50.5; DB 2; Length 27;
  Best Local Similarity 41.4%; Pred. No. 3.3e+02;
  Matches 12; Conservative 4; Mismatches 8; Indels 5; Gaps 1;

```
Qy        292 VPPTVQKPTTVNVPTTEVSPTSQKTTTKT 320
              |||| : |     || ||:| ::  |  |
Db          1 VPPTEEVP-----PTEEVTPPTEGETPPT 24
```

RESULT 5
US-08-450-360-12
; Sequence 12, Application US/08450360
; Patent No. 5656457
; GENERAL INFORMATION:
;   APPLICANT: Parkes, Deborah Lynn
;   APPLICANT: Coates, Stephen Ralph
;   TITLE OF INVENTION: Herpes Simplex Virus Type 2-Glycoprotein G
;   TITLE OF INVENTION: Proteins and Polypeptides
;   NUMBER OF SEQUENCES: 12
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Leona L. Lauder
;   STREET: 6 Mariposa Court
;   CITY: Tiburon
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94920
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/450,360
;       FILING DATE:  25-MAY-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/129,021
;       FILING DATE:  29-SEP-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0012.2A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-435-2034
;       TELEFAX:  415-435-0727
;   GENERAL INFORMATION:
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  26 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-450-360-12

  Query Match             2.3%;  Score 48;  DB 1;  Length 26;
  Best Local Similarity  41.7%;  Pred. No. 5.4e+02;
  Matches   10;  Conservative    2;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        311 PTSQKTTTKTTTPNAQATRSTPVS 334
              |||   | : |||  | |   | :
Db          1 PTSTHATPRPTTPGPQTTPPGPAT 24


RESULT 6
US-08-541-759B-4
; Sequence 4, Application US/08541759B
; Patent No. 5861160
;  GENERAL INFORMATION:
;    APPLICANT:  Quick, Douglas P.
;    APPLICANT:  Welter, Mark W.
;    APPLICANT:  Welter, Joseph
;    APPLICANT:  Welter, Lisa M.
;    TITLE OF INVENTION:  ISOSPORA SUIS VACCINE
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5861160el Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/541,759B
;      FILING DATE:  10-OCT-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klesner, Sharon N.
;      REGISTRATION NUMBER:  36,335
;      REFERENCE/DOCKET NUMBER:  Quick1a
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-541-759B-4

  Query Match             2.3%;  Score 47.5;  DB 2;  Length 27;
  Best Local Similarity  37.9%;  Pred. No. 6.3e+02;
  Matches   11;  Conservative    5;  Mismatches    8;  Indels    5;  Gaps    1;

Qy        292 VPPTVQKPTTVNVPTTEVSPTSQKTTTKT 320
              :||| : |     || ||:| ::  |  |
Db          1 LPPTEEVP-----PTEEVTPPTEGETPPT 24


RESULT 7
US-09-043-731-23
; Sequence 23, Application US/09043731A
; Patent No. 6344203
; GENERAL INFORMATION:
;  APPLICANT: The Austin Research Institute
;  TITLE OF INVENTION: Mimicking Peptides in Cancer Therapy
;  FILE REFERENCE: CALA-200
;  CURRENT APPLICATION NUMBER: US/09/043,731A
;  CURRENT FILING DATE: 1998-06-23
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: single
;   OTHER INFORMATION: stranded linear peptide
US-09-043-731-23

  Query Match             2.3%;  Score 47.5;  DB 4;  Length 29;
  Best Local Similarity  48.0%;  Pred. No. 7e+02;
  Matches   12;  Conservative    3;  Mismatches    9;  Indels    1;  Gaps    1;

Qy        299 PTTVNVPTTE-VSPTSQKTTTKTTT 322
              |||  : ||  |:||   | |:| |
Db          3 PTTTPISTTTMVTPTPTPTGTQTPT 27


RESULT 8
US-08-210-266A-4
; Sequence 4, Application US/08210266A
; Patent No. 5545619
;  GENERAL INFORMATION:
;    APPLICANT:  Atkinson, John P.
;    APPLICANT:  Hourcade, Dennis
;    APPLICANT:  Krych, Malgorzata
;    TITLE OF INVENTION:  Modified Complement System
;    TITLE OF INVENTION:  Regulators
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center, 1201 West Peachtree,
;      STREET:  Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,266A
;      FILING DATE:  18-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/695,514
;      FILING DATE:  03-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
```

; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)873-8794
; TELEFAX: (404)873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-210-266A-4

Query Match 2.3%; Score 47; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 SDPLPECR 221
||||||||
Db 1 SDPLPECR 8

1, 34

RESULT 9
US-08-688-675-4
; Sequence 4, Application US/08688675
; Patent No. 5719127
; GENERAL INFORMATION:
; APPLICANT: Atkinson, John P.
; APPLICANT: Hourcade, Dennis
; APPLICANT: Krych, Malgorzata
; TITLE OF INVENTION: Modified Complement System Regulators
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,675
; FILING DATE: 30-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/210,266
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/695,514
; FILING DATE: 03-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU101div2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)873-8794
; TELEFAX: (404)873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-688-675-4

Query Match 2.3%; Score 47; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 SDPLPECR 221
||||||||
Db 1 SDPLPECR 8

RESULT 10
US-08-477-860C-4
; Sequence 4, Application US/08477860C
; Patent No. 6010873
; GENERAL INFORMATION:
; APPLICANT: Atkinson, John P.
; APPLICANT: Hourcade, Dennis
; APPLICANT: Krych, Malgorzata
; TITLE OF INVENTION: Modified Complement System Regulators
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,860C
; FILING DATE: 7-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/210,266
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/695,514
; FILING DATE: 03-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU 101 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)873-8794
; TELEFAX: (404)873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-477-860C-4

Query Match 2.3%; Score 47; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 SDPLPECR 221
||||||||
Db 1 SDPLPECR 8

RESULT 11
US-08-169-948B-22
; Sequence 22, Application US/08169948B
; Patent No. 5861271
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; APPLICANT: Ward, Michael
; APPLICANT: Clarkson, Kathleen

```
;    APPLICANT:  Collier, Katherine
;    APPLICANT:  Larenas, Edmund
;    TITLE OF INVENTION:  No. 5861271el Cellulase Enzymes and Systems
;    TITLE OF INVENTION:  For Their Expression
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International
;      STREET:  180 Kimball Way
;      CITY:  South San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/169,948B
;      FILING DATE:  DEC 17 1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Horn, Margaret A.
;      REGISTRATION NUMBER:  33,401
;      REFERENCE/DOCKET NUMBER:  GC226
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 742-7536
;      TELEFAX:  (415)742-7217
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-169-948B-22

  Query Match             2.3%;  Score 47;  DB 2;  Length 27;
  Best Local Similarity   40.9%;  Pred. No. 7.1e+02;
  Matches    9;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        311 PTSQKTTTKTTTPNAQATRSTP 332
              |    :|| :||| :  | |:|
Db          1 PPPASSTTFSTTPRSSTTSSSP 22


RESULT 12
US-08-448-873-22
; Sequence 22, Application US/08448873
; Patent No. 5874276
;  GENERAL INFORMATION:
;    APPLICANT:  Fowler, Timothy
;    APPLICANT:  Ward, Michael
;    APPLICANT:  Clarkson, Kathleen
;    APPLICANT:  Collier, Katherine A.
;    APPLICANT:  Larenas, Edmund
;    TITLE OF INVENTION:  No. 5874276el Cellulase Enzymes and Systems
;    TITLE OF INVENTION:  For Their Expressions
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International
;      STREET:  180 Kimball Way
;      CITY:  South San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,873
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/169,948
;      FILING DATE:  17-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stone, Christopher L.
;      REGISTRATION NUMBER:  35,696
;      REFERENCE/DOCKET NUMBER:  GC226D14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 742-7555
;      TELEFAX:  (415)742-7217
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-448-873-22

  Query Match             2.3%;  Score 47;  DB 2;  Length 27;
  Best Local Similarity   40.9%;  Pred. No. 7.1e+02;
  Matches    9;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        311 PTSQKTTTKTTTPNAQATRSTP 332
              |    :|| :||| :  | |:|
Db          1 PPPASSTTFSTTPRSSTTSSSP 22


RESULT 13
US-08-382-452D-22
; Sequence 22, Application US/08382452D
; Patent No. 6268196
;  GENERAL INFORMATION:
;    APPLICANT:  Fowler, Timothy
;    APPLICANT:  Clarkson, Kathleen A.
;    APPLICANT:  Ward, Michael
;    APPLICANT:  Collier, Katherine D.
;    APPLICANT:  Larenas, Edmund A.
;    TITLE OF INVENTION:  NOVEL CELLULOSE ENZYMES AND SYSTEMS
;    TITLE OF INVENTION:  FOR THEIR EXPRESSION
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genencor International
;      STREET:  925 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/382,452D
;      FILING DATE:  February 1, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Christopher L. Stone
;      REGISTRATION NUMBER:  36,696
;      REFERENCE/DOCKET NUMBER:  GC226-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 742-7555
;      TELEFAX:  (415)742-7217
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-382-452D-22

  Query Match             2.3%;  Score 47;  DB 3;  Length 27;
  Best Local Similarity   40.9%;  Pred. No. 7.1e+02;
```

```
  Matches    9;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        311 PTSQKTTTKTTTPNAQATRSTP 332
              |    :|| :||| :  | |:|
Db          1 PPPASSTTFSTTPRSSTTSSSP 22


RESULT 14
US-08-507-362A-10
; Sequence 10, Application US/08507362A
; Patent No. 6562340
;   GENERAL INFORMATION:
;        APPLICANT: Bedford, Michael
;                   Morgan, Andrew
;                   Fowler, Timothy
;                   Ward, Michael
;                   Clarkson, Kathleen
;                   Collier, Katherine
;                   Larenas, Edmund
;        TITLE OF INVENTION: An Enzyme Feed Additive and Animal Feed Including It
;        NUMBER OF SEQUENCES: 21
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genencor International
;             STREET: 925 Page Mill Road
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/507,362A
;             FILING DATE: 27-Oct-1995
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Castaneda, Janet
;             REGISTRATION NUMBER: 33,228
;             REFERENCE/DOCKET NUMBER: GC226-3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650) 846-4072
;             TELEFAX: (650)845-6504
;   INFORMATION FOR SEQ ID NO: 10:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 27 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-08-507-362A-10

  Query Match             2.3%;  Score 47;  DB 4;  Length 27;
  Best Local Similarity   40.9%;  Pred. No. 7.1e+02;
  Matches    9;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        311 PTSQKTTTKTTTPNAQATRSTP 332
              |    :|| :||| :  | |:|
Db          1 PPPASSTTFSTTPRSSTTSSSP 22


RESULT 15
US-08-288-059-2
; Sequence 2, Application US/08288059
; Patent No. 5827666
;  GENERAL INFORMATION:
;    APPLICANT:  FINN, OLIVERA J.
;    APPLICANT:  FONTENOT, J. D.
;    APPLICANT:  MONTELARO, RONALD C.
;    TITLE OF INVENTION:  SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN
;    TITLE OF INVENTION:  AND MUCIN-LIKE PEPTIDES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,059
;      FILING DATE:  08-AUG-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CHAPIN, MARLANA K.
;      REGISTRATION NUMBER:  35,843
;      REFERENCE/DOCKET NUMBER:  61137/205204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861-3711
;      TELEFAX:  202-822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-288-059-2

  Query Match             2.3%;  Score 46.5;  DB 2;  Length 26;
  Best Local Similarity   50.0%;  Pred. No. 7.5e+02;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    1;  Gaps    1;

Qy        299 PTTVNV-PTTEVSPTSQKTTTKTT 321
              |||  :  || |:||   | | ||
Db          1 PTTTPITTTTTVTPTPTPTPTGTT 24


Search completed: November 18, 2003, 07:44:07
Job time : 22 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 18, 2003, 07:42:11 ; Search time 49 Seconds
(without alignments)
1234.180 Million cell updates/sec

Title: US-09-623-035-2
Perfect score: 2064
Sequence: 1 MTVARPSVPAALPLLGELPR.....HTCFTLTGLLGTLVTMGLLT 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 436344

Minimum DB seq length: 6
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	147	7.1	29	15	AAR50086	Decay accelerating
2	89.5	4.3	28	23	AAM50805	Human complement r
3	64.5	3.1	23	18	AAW31838	Peptide 4 from the
4	61	3.0	10	22	AAG97023	Human complementar
5	58	2.8	10	22	AAG97081	Human complementar
6	58	2.8	10	22	AAG97083	Human complementar
7	57	2.8	10	22	AAG97043	Human complementar
8	55	2.7	10	22	AAG97037	Human complementar
9	55	2.7	10	22	AAG97045	Human complementar
10	54	2.6	10	22	AAG97089	Human complementar
11	52.5	2.5	19	21	AAY65634	Oestrogen receptor
12	52.5	2.5	19	23	AAU86371	Oestrogen receptor
13	52	2.5	10	22	AAG96997	Human complementar
14	52	2.5	10	22	AAG97001	Human complementar
15	52	2.5	10	22	AAG97015	Human complementar
16	52	2.5	10	22	AAG97059	Human complementar
17	52	2.5	10	22	AAG97091	Human complementar
18	51.5	2.5	27	24	ABU07673	Human haptoglobin
19	51	2.5	10	22	AAG96999	Human complementar
20	51	2.5	10	22	AAG97003	Human complementar
21	51	2.5	10	22	AAG97005	Human complementar
22	51	2.5	10	22	AAG97007	Human complementar
23	51	2.5	10	22	AAG97009	Human complementar
24	51	2.5	10	22	AAG97011	Human complementar
25	51	2.5	10	22	AAG97013	Human complementar
26	51	2.5	10	22	AAG97017	Human complementar
27	51	2.5	10	22	AAG97031	Human complementar
28	51	2.5	10	22	AAG97033	Human complementar
29	51	2.5	10	22	AAG97041	Human complementar
30	51	2.5	10	22	AAG97049	Human complementar
31	51	2.5	10	22	AAG97051	Human complementar
32	51	2.5	10	22	AAG97053	Human complementar
33	51	2.5	10	22	AAG97055	Human complementar
34	51	2.5	10	22	AAG97069	Human complementar
35	51	2.5	10	22	AAG97071	Human complementar
36	51	2.5	10	22	AAG97073	Human complementar
37	51	2.5	10	22	AAG97075	Human complementar
38	51	2.5	10	22	AAG97077	Human complementar
39	51	2.5	10	22	AAG97079	Human complementar
40	51	2.5	10	22	AAG97085	Human complementar
41	51	2.5	10	22	AAG97093	Human complementar
42	51	2.5	10	22	AAG97095	Human complementar
43	51	2.5	10	22	AAG97097	Human complementar
44	51	2.5	10	22	AAG97105	Human complementar
45	51	2.5	10	22	AAG97107	Human complementar

ALIGNMENTS

RESULT 1
AAR50086
ID AAR50086 standard; Protein; 29 AA.
XX
AC AAR50086;
XX
DT 25-MAR-2003 (updated)
DT 27-OCT-1994 (first entry)
XX
DE Decay accelerating factor derived attachment signal.
XX
KW Fusion protein; xenograft; transplantation; membrane; transplant;
KW membrane cofactor protein; MCP; decay accelerating factor; DAF; GPI;
KW glycophosphatidyl inositol; anchor; intermembrane transfer.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 13..29
FT /note= "C-terminal hydrophobic domain."
XX
PN WO9406903-A1.
XX
PD 31-MAR-1994.
XX
PF 22-SEP-1993; 93WO-US08889.
XX
PR 22-SEP-1992; 92US-0948521.
XX
PA (DNXB-) DNX BIOTHERAPEUTICS INC.
XX
PI Byrne GW, Kooyman DL, Logan JS;

XX
DR WPI; 1994-118454/14.
XX
PT Delivery of proteins, using mobile cell, by intermembrane
PT transfer - for pre-accommodation of xenogeneic organ transplants
XX
PS Disclosure; Page 43; 104pp; English.
XX
CC A protein of interest may be delivered to the tissue of a target
CC animal by providing mobile cells which bear the protein on their
CC membranes, linked to a glycophosphatidyl inositol (GPI) anchor. The
CC cells circulate in the target animal, coming into contact with the
CC tissue and transfer the protein to the tissue by intermembrane
CC transfer. Intermembrane transfer by this method occurs at a greater
CC rate than naturally in the animal. This sequence is an attachment
CC signal derived from decay accelerating factor (DAF). DAF is
CC naturally GPI linked. Preferably, the protein of interest is
CC expressed as a fusion product, the fusion protein comprising this
CC sequence and the protein (or fragment) of interest. The method is
CC especially useful for modifying pig organs for xenotransplantation
CC into humans. See AAQ58894.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 29 AA;

Query Match 7.1%; Score 147; DB 15; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        353 SGTTRLLSGHTCFTLTGLLGTLVTMGLLT 381
              |||||||||||||||||||||||||||||
Db          1 SGTTRLLSGHTCFTLTGLLGTLVTMGLLT 29
```

1, 11, 13, 34

RESULT 2
AAM50805
ID AAM50805 standard; Peptide; 28 AA.
XX
AC AAM50805;
XX
DT 01-MAY-2002 (first entry)
XX
DE Human complement receptor type 1 tryptic peptide 54d.
XX
KW Complement receptor type 1; CR1; C3b/C4b receptor; human;
KW inflammation; myocardial infarction; reperfusion injury;
KW antiinflammatory; therapy; diagnosis; receptor; purification.
XX
OS Homo sapiens.
XX
PN US6316604-B1.
XX
PD 13-NOV-2001.
XX
PF 05-JUN-1995; 95US-0463959.
XX
PR 03-APR-1989; 89US-0332865.
PR 06-DEC-1994; 94US-0350238.
PR 24-FEB-1993; 93US-0026134.
PR 01-APR-1988; 88US-0176532.
XX
PA (AVAN-) AVANT IMMUNOTHERAPEUTICS INC.
XX
PI Fearon DT, Klickstein LB, Wong WW, Carson GR, Concino MF, Ip SH;
PI Makrides SC, Marsh HC;
XX
DR WPI; 2002-121028/16.
XX
PT Purifying a recombinant soluble complement receptor type 1 molecule or
PT its fragment, useful for diagnosing or treating disorders associated
PT with complement activity, using cation exchange chromatography provides
PT molecules -
XX
PS Example; Column 25; 44pp; English.
XX
CC The present sequence is that of tryptic peptide 54d of human
CC erythrocyte C3b/C4b receptor CR1 (complement receptor type 1).
CC The peptide corresponds to amino acids 204-231 and 654-681 of a
CC full-length CR1 sequence (see AAM50797) predicted from isolated cDNA
CC clones. The invention provides recombinant, soluble CR1 polypeptides
CC that lack the transmembrane domain of the native protein, and methods
CC of purifying them. CR1 polypeptides and nucleic acids have use in
CC the diagnosis and therapy of disorders involving complement activity
CC and various immune system or inflammatory disorders. Secreted,
CC recombinant CR1 has been produced in large quantities using CHO host
CC cells grown in serum-free medium. The purified protein was used to
CC reduce damage caused by inflammation, and to reduce myocardial
CC infarct size and prevent reperfusion injury in rats.
XX
SQ Sequence 28 AA;

Query Match 4.3%; Score 89.5; DB 23; Length 28;
Best Local Similarity 59.3%; Pred. No. 3.9;
Matches 16; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

```
Qy        257 FTMIGEHSIYCTVNNDE-GEWSGPPPE 282
              | ::|| ||||| |:|: | |||| |:
Db          2 FELVGEPSIYCTSNDDQVGIWSGPAPQ 28
```

RESULT 3
AAW31838
ID AAW31838 standard; peptide; 23 AA.
XX
AC AAW31838;
XX
DT 26-MAR-1998 (first entry)
XX
DE Peptide 4 from the short consensus repeat 3 of complement receptor 1.
XX
KW Short consensus repeat 3; long homologous repeat A; LHR-A;
KW complement receptor 1; CR1; complement inhibition;
KW anti-haemolytic activity; inflammation; thrombotic condition;
KW inappropriate complement activation; ARDS; Alzheimer's disease.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9731944-A1.
XX
PD 04-SEP-1997.
XX
PF 26-FEB-1997; 97WO-EP00994.
XX
PR 02-MAR-1996; 96GB-0004518.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Edge CM, Mossakowska DEI, Smith RAG;
XX
DR WPI; 1997-448630/41.
XX
PT Peptide derived from short consensus repeat 3 of human complement
PT receptor 1 - between amino acids Cys154-Gly186, useful to treat
PT disorder or disease associated with inflammation or inappropriate
PT complement activation
XX
PS Claim 13; Page 25; 33pp; English.
XX
CC Synthetic soluble peptides AAW31835-38 are derived from the short
CC consensus repeat 3 of the long homologous repeat A (LHR-A) of human
CC complement receptor 1 (CR1). The present peptide is located between
CC residues 164 and 186. These peptides have functional complement
CC inhibitory, including anti-haemolytic, activity. Enhanced activity
CC may be achieved by linking the peptides to a core structure
CC (e.g. the MAP peptide) to produce a multimeric or chimeric polypeptide.

```
CC   This polypeptide can be used to treat a disorder or disease associated
CC   with inflammation or inappropriate complement activation. It can also be
CC   used to treat a thrombotic condition, adult respiratory distress
CC   syndrome (ARDS), wounds, Alzheimer's disease or a CNS inflammatory
CC   disorder, or delay hyperacute allograft or xenograft rejection.
XX
SQ   Sequence   23 AA;

  Query Match             3.1%;  Score 64.5;  DB 18;  Length 23;
  Best Local Similarity   56.5%;  Pred. No. 2.6e+02;
  Matches   13;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        257 FTMIGEHSIYCTVNNDE-GEWSG 278
              | ::|| ||| | |:|: | |||
Db          1 FELVGEPSIYSTSNDDQVGIWSG 23


RESULT 4
AAG97023
ID   AAG97023 standard; Peptide; 10 AA.
XX
AC   AAG97023;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Human complementary peptide, SEQ ID NO: 3217.
XX
KW   Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS   Homo sapiens.
XX
PN   WO200142277-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB04776.
XX
PR   13-DEC-1999;   99GB-0029464.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-408419/43.
XX
PT   A set of peptide ligands consisting of specific complementary peptides
PT   to proteins encoded by genes of the human genome, useful in an assay
PT   for screening and identifying of one or more novel peptides which are
PT   drug candidates or pro-drugs -
XX
PS   Example 4; Page 506; 646pp; English.
XX
CC   The invention relates to a set of complementary peptide ligands
CC   generated from the human genome. The complementary peptides
CC   interact with their relevant target proteins encoded in the human
CC   genome. They can be used as reagents in drug discovery and as lead
CC   ligands to facilitate drug design and development. The present
CC   sequence is a complementary peptide provided in the specification.
XX
SQ   Sequence   10 AA;

  Query Match             3.0%;  Score 61;  DB 22;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        278 GPPPECRGKS 287
              ||||||||||
Db          1 GPPPECRGKS 10


RESULT 5
AAG97081
ID   AAG97081 standard; Peptide; 10 AA.
XX
AC   AAG97081;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Human complementary peptide, SEQ ID NO: 3275.
XX
KW   Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS   Homo sapiens.
XX
PN   WO200142277-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB04776.
XX
PR   13-DEC-1999;   99GB-0029464.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-408419/43.
XX
PT   A set of peptide ligands consisting of specific complementary peptides
PT   to proteins encoded by genes of the human genome, useful in an assay
PT   for screening and identifying of one or more novel peptides which are
PT   drug candidates or pro-drugs -
XX
PS   Example 4; Page 514; 646pp; English.
XX
CC   The invention relates to a set of complementary peptide ligands
CC   generated from the human genome. The complementary peptides
CC   interact with their relevant target proteins encoded in the human
CC   genome. They can be used as reagents in drug discovery and as lead
CC   ligands to facilitate drug design and development. The present
CC   sequence is a complementary peptide provided in the specification.
XX
SQ   Sequence   10 AA;

  Query Match             2.8%;  Score 58;  DB 22;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        151 WSTAVEFCKK 160
              ||||||||||
Db          1 WSTAVEFCKK 10


RESULT 6
AAG97083
ID   AAG97083 standard; Peptide; 10 AA.
XX
AC   AAG97083;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Human complementary peptide, SEQ ID NO: 3277.
XX
KW   Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS   Homo sapiens.
XX
PN   WO200142277-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB04776.
XX
PR   13-DEC-1999;   99GB-0029464.
XX
PA   (PROT-) PROTEOM LTD.
```

```
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-408419/43.
XX
PT   A set of peptide ligands consisting of specific complementary peptides
PT   to proteins encoded by genes of the human genome, useful in an assay
PT   for screening and identifying of one or more novel peptides which are
PT   drug candidates or pro-drugs -
XX
PS   Example 4; Page 514; 646pp; English.
XX
CC   The invention relates to a set of complementary peptide ligands
CC   generated from the human genome. The complementary peptides
CC   interact with their relevant target proteins encoded in the human
CC   genome. They can be used as reagents in drug discovery and as lead
CC   ligands to facilitate drug design and development. The present
CC   sequence is a complementary peptide provided in the specification.
XX
SQ   Sequence   10 AA;

  Query Match             2.8%;  Score 58;  DB 22;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        361 GHTCFTLTGL 370
              ||||||||||
Db          1 GHTCFTLTGL 10


RESULT 7
AAG97043
ID   AAG97043 standard; Peptide; 10 AA.
XX
AC   AAG97043;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Human complementary peptide, SEQ ID NO: 3237.
XX
KW   Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS   Homo sapiens.
XX
PN   WO200142277-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB04776.
XX
PR   13-DEC-1999;   99GB-0029464.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-408419/43.
XX
PT   A set of peptide ligands consisting of specific complementary peptides
PT   to proteins encoded by genes of the human genome, useful in an assay
PT   for screening and identifying of one or more novel peptides which are
PT   drug candidates or pro-drugs -
XX
PS   Example 4; Page 508; 646pp; English.
XX
CC   The invention relates to a set of complementary peptide ligands
CC   generated from the human genome. The complementary peptides
CC   interact with their relevant target proteins encoded in the human
CC   genome. They can be used as reagents in drug discovery and as lead
CC   ligands to facilitate drug design and development. The present
CC   sequence is a complementary peptide provided in the specification.
XX
SQ   Sequence   10 AA;

  Query Match             2.8%;  Score 57;  DB 22;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        123 TVVEYECRPG 132
              ||||||||||
Db          1 TVVEYECRPG 10


RESULT 8
AAG97037
ID   AAG97037 standard; Peptide; 10 AA.
XX
AC   AAG97037;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Human complementary peptide, SEQ ID NO: 3231.
XX
KW   Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS   Homo sapiens.
XX
PN   WO200142277-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB04776.
XX
PR   13-DEC-1999;   99GB-0029464.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-408419/43.
XX
PT   A set of peptide ligands consisting of specific complementary peptides
PT   to proteins encoded by genes of the human genome, useful in an assay
PT   for screening and identifying of one or more novel peptides which are
PT   drug candidates or pro-drugs -
XX
PS   Example 4; Page 508; 646pp; English.
XX
CC   The invention relates to a set of complementary peptide ligands
CC   generated from the human genome. The complementary peptides
CC   interact with their relevant target proteins encoded in the human
CC   genome. They can be used as reagents in drug discovery and as lead
CC   ligands to facilitate drug design and development. The present
CC   sequence is a complementary peptide provided in the specification.
XX
SQ   Sequence   10 AA;

  Query Match             2.7%;  Score 55;  DB 22;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        247 QSVTYACNKG 256
              ||||||||||
Db          1 QSVTYACNKG 10


RESULT 9
AAG97045
ID   AAG97045 standard; Peptide; 10 AA.
XX
AC   AAG97045;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Human complementary peptide, SEQ ID NO: 3239.
XX
```

KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -
XX
PS Example 4; Page 509; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match 2.7%; Score 55; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 358 LLSGHTCFTL 367
||||||||||
Db 1 LLSGHTCFTL 10

RESULT 10
AAG97089
ID AAG97089 standard; Peptide; 10 AA.
XX
AC AAG97089;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 3283.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -
XX
PS Example 4; Page 515; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match 2.6%; Score 54; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 291 KVPPTVQKPT 300
||||||||||
Db 1 KVPPTVQKPT 10

RESULT 11
AAY65634
ID AAY65634 standard; Peptide; 19 AA.
XX
AC AAY65634;
XX
DT 01-FEB-2000 (first entry)
XX
DE Oestrogen receptor beta ERE binding peptide 29E-beta.
XX
KW Oestrogen receptor; estrogen; estradiol; oestrogen response element;
KW ERE; binding; biological activity; fingerprint; molecular braille;
KW cellular braille; modulation; tamoxifen; breast cancer; ovarian cancer;
KW menopause; osteoporosis; selective oestrogen receptor modulator;
KW identification; characterisation; classification.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9954728-A2.
XX
PD 28-OCT-1999.
XX
PF 26-MAR-1999; 99WO-US06664.
XX
PR 23-APR-1998; 98US-0082756.
PR 09-SEP-1998; 98US-0099656.
PR 08-JAN-1999; 99US-0115345.
XX
PA (NOVA-) NOVALON PHARM CORP.
XX
PI Paige LA, Hamilton PT, Fowlkes DM, Buehrer B, Barnett T;
PI McDonnell DP, Christensen DJ;
XX
DR WPI; 2000-013281/01.
XX
PT Methods for identifying new receptor modulators, especially estrogen
PT modulators to treat tamoxifen refractory breast cancer -
XX
PS Example 2.2; Page 165; 219pp; English.
XX
CC The present invention describes a method for predicting the biological
CC activity of new receptor modulating compounds (I) using novel oligomeric
CC peptides (biokeys) which have differential abilities to bind to 2
CC different receptor conformations. The method is used to identify new
CC drugs that are physiological or pharmacological agonists/antagonists and
CC that target various receptors, which are involved in certain disease
CC conditions. The system may be used as a primary screening tool to
CC identify hits, to classify lead compounds from a drug screen to,
CC characterise selective oestrogen receptor modulators (SERMs) in terms of

CC agonist and antagonist function and to predict possible clinical effects
CC of SERMs such as tissue and receptor specificity. The method can also be
CC applied to the fractionation of mixtures of SERMs to determine which
CC components are producing agonistic and antagonistic activity. The method
CC may be used with other receptors (e.g. progesterone, androgen,
CC glucocorticoid, thyroid, vitamin D, beta-adrenergic, dopamine and
CC epidermal growth factor, to identify, characterise and classify
CC modulators of receptor activity. Peptides comprising a LXXLL motif may
CC be used to modulate the oestrogen receptor in treating e.g. breast and
CC ovarian cancer and ameliorating the effects of menopause, including
CC osteoporosis. AAY65439 to AAY65652 represent oestrogen receptor,
CC estradiol receptor and oestrogen response element binding peptides given
CC in the exemplification of the present invention. AAZ35740 to AAZ35745
CC represent oligonucleotides used in the exemplification of the present
CC invention.
XX
SQ Sequence 19 AA;

Query Match 2.5%; Score 52.5; DB 21; Length 19;
Best Local Similarity 47.6%; Pred. No. 1.7e+03;
Matches 10; Conservative 1; Mismatches 5; Indels 5; Gaps 1;

Qy 209 SSVQWSDPLPECREIYCPAPP 229
|| | :| ||||||
Db 1 SSYDW-----QCPSWYCPAPP 16

RESULT 12
AAU86371
ID AAU86371 standard; Peptide; 19 AA.
XX
AC AAU86371;
XX
DT 21-MAY-2002 (first entry)
XX
DE Oestrogen receptor beta binding peptide 29E-beta.
XX
KW Oestrogen receptor; breast cancer; combinatorial peptide library;
KW receptor modulating compound.
XX
OS Synthetic.
XX
PN WO200204956-A2.
XX
PD 17-JAN-2002.
XX
PF 11-JUL-2001; 2001WO-US21867.
XX
PR 12-JUL-2000; 2000US-0614865.
PR 21-MAY-2001; 2001US-0860688.
XX
PA (KARO-) KARO BIO USA INC.
XX
PI Fowlkes DM, Barnett TR, Buehrer B;
XX
DR WPI; 2002-154969/20.
XX
PT Identifying receptor-binding peptides comprises screening combinatorial
PT peptide library presented in form of cells each of which coexpress one
PT peptide member and receptor with signal producing system for reporting
PT binding
XX
PS Disclosure; Page 147; 175pp; English.
XX
CC The invention relates to identifying a binding peptide which binds a
CC receptor and which is a member of a combinatorial library of peptides,
CC comprising screening a combinatorial peptide library presented in the
CC form of cells which coexpress the receptor or its ligand-binding receptor
CC moiety and one member of the library, together with a signal producing
CC system for reporting binding of the peptide to the receptor. Also
CC included is a method for predicting the receptor-modulating activity of a
CC compound which modulates the biological activity of a receptor
CC comprising (a) identifying peptides which bind the receptor by the
CC method above, (b) using a number of the peptides to predict the receptor-
CC modulating activity of a compound by (i) providing a panel of
CC identified peptides, where the members differ in their ability to bind
CC to the receptor depending on reference conformations the receptor is
CC in, where the effect of a number of reference substances known to
CC modulate the biological activity of the receptor on the binding of each
CC member of the panel is known and is characterised as a reference
CC fingerprint for each reference substance, (ii) screening a test substance
CC of unknown activity relative to the receptor to determine its effect on
CC the binding of each member of the panel to the receptor, thereby
CC obtaining a test fingerprint for the test substance, (iii) comparing the
CC test fingerprint to the reference fingerprints and (iv) predicting the
CC biological activity of the test substance based on the assumption that
CC its biological activity will be similar to that of reference substances
CC with similar fingerprints. The method is useful for identifying a binding
CC peptide which binds a vertebrate, mammalian, preferably human receptor,
CC an intracellular, nuclear, oestrogen or androgen receptor. The identified
CC peptides which bind to the receptor are useful for predicting the
CC receptor-modulating activity of a compound (e.g. ant/agonists).
CC The receptor-binding library members are useful in the prediction of the
CC ability of small organic molecules, suitable for pharmaceutical use
CC (e.g. in the case of oestrogen receptors, for breast cancer treatment),
CC to interact with the receptor. The analyte-binding molecules can also be
CC used for in vivo imaging. The method has several advantages over whole
CC animal-based assay systems in that the same technology can be applied to
CC a variety of different receptors, the system can be used for high
CC throughput screening and compound characterisation, and gives very
CC distinct patterns for agonists and antagonists of receptor activity using
CC very much less protein. The present sequence is an oestrogen receptor
CC binding peptide from a combinatorial peptide library.
XX
SQ Sequence 19 AA;

Query Match 2.5%; Score 52.5; DB 23; Length 19;
Best Local Similarity 47.6%; Pred. No. 1.7e+03;
Matches 10; Conservative 1; Mismatches 5; Indels 5; Gaps 1;

Qy 209 SSVQWSDPLPECREIYCPAPP 229
|| | :| ||||||
Db 1 SSYDW-----QCPSWYCPAPP 16

RESULT 13
AAG96997
ID AAG96997 standard; Peptide; 10 AA.
XX
AC AAG96997;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 3191.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are

PT drug candidates or pro-drugs -
XX
PS Example 4; Page 502; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match 2.5%; Score 52; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 196 LFGSTSSFCL 205
||||||||||
Db 1 LFGSTSSFCL 10

RESULT 14
AAG97001
ID AAG97001 standard; Peptide; 10 AA.
XX
AC AAG97001;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 3195.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -
XX
PS Example 4; Page 503; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match 2.5%; Score 52; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 316 TTTKTTTPNA 325
||||||||||
Db 1 TTTKTTTPNA 10

RESULT 15
AAG97015
ID AAG97015 standard; Peptide; 10 AA.
XX
AC AAG97015;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 3209.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -
XX
PS Example 4; Page 504; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match 2.5%; Score 52; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 192 TGYKLFGSTS 201
||||||||||
Db 1 TGYKLFGSTS 10

Search completed: November 18, 2003, 07:43:38
Job time : 50 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 18, 2003, 07:42:46 ; Search time 26 Seconds
(without alignments)
1409.241 Million cell updates/sec

Title: US-09-623-035-2
Perfect score: 2064
Sequence: 1 MTVARPSVPAALPLLGELPR.....HTCFTLTGLLGTLVTMGLLT 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 6074

Minimum DB seq length: 6
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	2.3	25	2	A48810	fibrinogen B beta
2	41.5	2.0	28	2	I52394	fibronectin, splic
3	40	1.9	26	2	S68902	xanthine dehydroge
4	40	1.9	27	1	TZAZ3	toxin III - snake-
5	40	1.9	29	2	I78537	copper transportin
6	39	1.9	28	2	I48349	fibronectin - mous
7	38	1.8	24	2	A84023	hypothetical prote
8	38	1.8	26	2	D32248	Ig kappa chain V r
9	38	1.8	26	2	A44036	collagen alpha 1(X
10	38	1.8	27	1	PGFF1	paragonial peptide
11	37.5	1.8	26	2	S65604	sec c 1 protein -
12	37	1.8	11	2	D45900	complement C3b rec
13	37	1.8	20	2	PC1151	equinatoxin 1C - s
14	37	1.8	20	2	B34016	tenebrosin B - sea
15	37	1.8	28	2	B64669	hypothetical prote
16	36.5	1.8	30	2	S59482	hydroxyproline-ric
17	36	1.7	20	2	PC1150	equinatoxin 1B - s
18	36	1.7	26	2	A61195	reverse transcript
19	36	1.7	27	2	PL0029	plasma protein Po2
20	36	1.7	27	2	A84477	5S ribosomal RNA [
21	36	1.7	28	2	E49533	T-cell receptor be
22	36	1.7	30	2	A27103	aspartate transami
23	35.5	1.7	29	2	C61384	trachael mucin gly
24	35	1.7	18	2	JU0125	polyphemusin II -
25	35	1.7	20	2	A34016	tenebrosin A - sea
26	35	1.7	21	2	S08590	NADH2 dehydrogenas
27	35	1.7	24	2	B49480	major immunophilin
28	35	1.7	25	2	I60083	glycophorin A - hu
29	34.5	1.7	20	2	B41299	T-cell receptor al
30	34.5	1.7	27	2	A47295	homeodomain protei
31	34.5	1.7	27	2	S28940	cyclic nucleotide-
32	34.5	1.7	29	2	A60604	glutathione peroxi
33	34	1.6	11	2	I33098	173K exoantigen -
34	34	1.6	15	2	PT0222	Ig heavy chain CDR
35	34	1.6	16	2	S38292	30K allergen - rye
36	34	1.6	18	2	I52614	u-plasminogen acti
37	34	1.6	19	2	PH1339	Ig heavy chain DJ
38	34	1.6	19	2	B28457	proteoglycan II, b
39	34	1.6	20	2	A61276	superoxide dismuta
40	34	1.6	21	2	I65270	collagen alpha 1(I
41	34	1.6	23	2	E39855	paralytic peptide
42	34	1.6	23	2	C39855	paralytic peptide
43	34	1.6	23	2	D39855	paralytic peptide
44	34	1.6	23	2	B60691	phycobilisome 29K
45	34	1.6	26	1	NTSR3L	neurotoxin III - E

ALIGNMENTS

RESULT 1
A48810
fibrinogen B beta subunit - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A48810
R;Roberts, L.R.; Nichols, L.A.; Holland, L.J.
Biochemistry 32, 11627-11637, 1993
A;Title: Transcriptional regulation of the Xenopus laevis B beta fibrinogen subunit gene
A;Reference number: A48810; MUID:94032285; PMID:8218230
A;Accession: A48810
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-25 <ROB>
A;Cross-references: GB:U05035; GB:S66373; NID:g450950; PIDN:AAA60463.1; PID:g450951
A;Note: sequence extracted from NCBI backbone (NCBIN:138880, NCBIP:138881)

Query Match 2.3%; Score 48; DB 2; Length 25;
Best Local Similarity 64.3%; Pred. No. 1.9e+03;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy         20 RLLLLVLLCLPAVW 33
              |:|||  ||: |||
Db          2 RVLLLFALCVSAVW 15
```

RESULT 2
I52394
fibronectin, splice form II - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C;Accession: I52394; I53217
R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different mes
A;Reference number: I52394; MUID:87026578; PMID:3021206
A;Accession: I52394
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-28 <RES>
A;Cross-references: GB:M14059; NID:g182700; PIDN:AAA52463.1; PID:g182703
R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 3, 221-226, 1984
A;Title: Human fibronectin: Molecular cloning evidence for two mRNA species differing by
A;Reference number: I53217; MUID:84158533; PMID:6200322
A;Accession: I53217
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 3-16 <RE2>
A;Cross-references: GB:K02273; NID:g182698; PIDN:AAA35850.1; PID:g182699
C;Genetics:
A;Gene: GDB:FN1

A;Cross-references: GDB:119135; OMIM:135600
A;Map position: 2q34-2q34
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C;Keywords: alternative splicing

Query Match 2.0%; Score 41.5; DB 2; Length 28;
Best Local Similarity 48.0%; Pred. No. 6.1e+03;
Matches 12; Conservative 3; Mismatches 5; Indels 5; Gaps 2;

```
Qy        294 PTVQKP-TTVNVPT----TEVSPTS 313
              | ||   ||:  ||    |:|:|||
Db          2 PLVQTAVTTIPAPTDLKFTQVTPTS 26
```

RESULT 3
S68902
xanthine dehydrogenase (EC 1.1.1.204) 18.4K chain - Veillonella atypica (fragment)
C;Species: Veillonella atypica
C;Date: 23-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
C;Accession: S68902
R;Gremer, L.; Meyer, O.
Eur. J. Biochem. 238, 862-866, 1996
A;Title: Characterization of xanthine dehydrogenase from the anaerobic bacterium Veillon
A;Reference number: S68900; MUID:96300255; PMID:8706691
A;Accession: S68902
A;Molecule type: protein
A;Residues: 1-26 <GRE>
A;Experimental source: DSM 1399
C;Keywords: 2Fe-2S; FAD; flavoprotein; heterotrimer; iron-sulfur protein; metalloprotein

Query Match 1.9%; Score 40; DB 2; Length 26;
Best Local Similarity 56.2%; Pred. No. 7.1e+03;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Qy        298 KPTTVNVPTTEVSPTS 313
              |  |||||| |:  |:
Db         10 KNVTVNVPTDEMLLTT 25
```

RESULT 4
TZAZ3
toxin III - snake-locks sea anemone
C;Species: Anemonia sulcata (snake-locks sea anemone)
C;Date: 30-Apr-1979 #sequence_revision 24-Sep-1981 #text_change 04-Oct-1996
C;Accession: A91446; A91674; A01798
R;Martinez, G.; Kopeyan, C.; Schweitz, H.; Lazdunski, M.
FEBS Lett. 84, 247-252, 1977
A;Title: Toxin III from Anemonia sulcata: primary structure.
A;Reference number: A91446; MUID:78084776; PMID:23311
A;Accession: A91446
A;Molecule type: protein
A;Residues: 1-27 <MAR>
R;Beress, L.; Wunderer, G.; Wachter, E.
Hoppe-Seyler's Z. Physiol. Chem. 358, 985-988, 1977
A;Title: Amino acid sequence of toxin III from Anemonia sulcata.
A;Reference number: A91674; MUID:78044787; PMID:21843
A;Accession: A91674
A;Molecule type: protein
A;Residues: 1-21,'SC',24-27 <BER>
C;Comment: Three disulfide bonds are present.
C;Superfamily: toxin III
C;Keywords: venom

Query Match 1.9%; Score 40; DB 1; Length 27;
Best Local Similarity 26.7%; Pred. No. 7.4e+03;
Matches 8; Conservative 2; Mismatches 8; Indels 12; Gaps 1;

```
Qy         28 CLPAVWGDCGLPPDVPNAQPALEGRTSFPE 57
              | |  || |               |:  :||
Db          3 CCPCYWGGCPW-----------GQNCYPE 20
```

RESULT 5
I78537
copper transporting P-type ATPase - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C;Accession: I78537
R;Thomas, G.R.; Forbes, J.R.; Roberts, E.A.; Walshe, J.M.; Cox, D.W.
Nature Genet. 9, 210-217, 1995
A;Title: The Wilson disease gene: spectrum of mutations and their consequences.
A;Reference number: I58128; MUID:95235569; PMID:7626145
A;Accession: I78537
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-29 <RES>
A;Cross-references: GB:S77450; NID:g957354; PIDN:AAB34087.1; PID:g957355
C;Genetics:
A;Gene: GDB:ATP7B
A;Cross-references: GDB:120494; OMIM:277900
A;Map position: 13q14.3-13q21.1

Query Match 1.9%; Score 40; DB 2; Length 29;
Best Local Similarity 31.8%; Pred. No. 8.1e+03;
Matches 7; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

```
Qy         28 CLPAVWGDCGLPPDVPNAQPAL 49
              |||:   :   | |:    | |
Db          8 CLPSALSEIQFPADIRQKSPPL 29
```

RESULT 6
I48349
fibronectin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C;Accession: I48349; S33445
R;Polly, P.; Nicholson, R.C.
Gene 137, 353-354, 1993
A;Title: Sequence of the mouse fibronectin-encoding gene promoter region.
A;Reference number: I48349; MUID:94131313; PMID:8299972
A;Accession: I48349
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-28 <RES>
A;Cross-references: EMBL:Z22729; NID:g297911; PIDN:CAA80422.1; PID:g297912
R;Polly, P.; Nicholson, R.C.
submitted to the EMBL Data Library, May 1993
A;Description: Nucleotide sequence of the murine fibronectin gene promoter region.
A;Reference number: S33445
A;Accession: S33445
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-27 <POL>
A;Cross-references: EMBL:Z22729
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repea

Query Match 1.9%; Score 39; DB 2; Length 28;
Best Local Similarity 80.0%; Pred. No. 9.1e+03;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy         20 RLLLLVLLCL 29
              ||||| :|||
Db          9 RLLLLAVLCL 18
```

RESULT 7
A84023
hypothetical protein BH2985 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: A84023
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiram
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132

```
A;Accession: A84023
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-24 <STO>
A;Cross-references: GB:AP001517; GB:BA000004; NID:g10175500; PIDN:BAB06704.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2985

  Query Match             1.8%;  Score 38;  DB 2;  Length 24;
  Best Local Similarity   35.3%;  Pred. No. 8.9e+03;
  Matches   12;  Conservative    5;  Mismatches    7;  Indels   10;  Gaps    2;

Qy        104 LNSASLKQPYITQNYFPVGTVVEYECRPGYRREP 137
              :|||::|:   |: |           | | ||||
Db          1 MNSAAIKKQ--TEKY--------GASRRGLRREP 24


RESULT 8
D32248
Ig kappa chain V region (5-27) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 30-May-1997
C;Accession: D32248
R;Bedzyk, W.D.; Johnson, L.S.; Riordan, G.S.; Voss Jr., E.W.
J. Biol. Chem. 264, 1565-1569, 1989
A;Title: Comparison of variable region primary structures within an anti-fluorescein idi
A;Reference number: A31485; MUID:89109167; PMID:2492278
A;Accession: D32248
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-26 <BED>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

  Query Match             1.8%;  Score 38;  DB 2;  Length 26;
  Best Local Similarity   58.3%;  Pred. No. 9.8e+03;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MTVARPSVPAAL 12
              ||  :||:||:|
Db          4 MTQTKPSLPASL 15


RESULT 9
A44036
collagen alpha 1(XII) chain - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 23-May-1997
C;Accession: A44036
R;Lunstrum, G.P.; McDonough, A.M.; Marinkovich, M.P.; Keene, D.R.; Morris, N.P.; Burgesc
J. Biol. Chem. 267, 20087-20092, 1992
A;Title: Identification and partial purification of a large, variant form of type XII cc
A;Reference number: A44036; MUID:93015874; PMID:1400326
A;Accession: A44036
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-26 <LUN>
A;Experimental source: skin
A;Note: sequence extracted from NCBI backbone (NCBIP:115795)
C;Superfamily: collagen alpha 1(XII) chain; fibronectin type III repeat homology; von Wi

  Query Match             1.8%;  Score 38;  DB 2;  Length 26;
  Best Local Similarity   41.2%;  Pred. No. 9.8e+03;
  Matches    7;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          57 EDTVITYKCEESFVKIP 73
              || :||: || :    |
Db           2 EDNLITFVCETATSSCP 18


RESULT 10
PGFF1
paragonial peptide PS-1 - fruit fly (Drosophila funebris)
C;Species: Drosophila funebris
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 21-Feb-1997
C;Accession: A01643
R;Baumann, H.; Wilson, K.J.; Chen, P.S.; Humbel, R.E.
Eur. J. Biochem. 52, 521-529, 1975
A;Title: The amino-acid sequence of a peptide (PS-1) from Drosophila funebris: a paragon
A;Reference number: A01643; MUID:77245908; PMID:1236144
A;Accession: A01643
A;Molecule type: protein
A;Residues: 1-27 <BAU>
A;Note: 2-Leu was found in 30% of the molecules
C;Comment: This peptide, produced in the paragonial (accessory sex) gland of adult male f
C;Genetics:
A;Gene: FlyBase:Dfun/PapC
A;Cross-references: FlyBase:FBgn0004112
C;Superfamily: paragonial peptide PS-1

  Query Match             1.8%;  Score 38;  DB 1;  Length 27;
  Best Local Similarity   26.9%;  Pred. No. 1e+04;
  Matches    7;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        303 NVPTTEVSPTSQKTTTKTTTPNAQAT 328
              :||:   :  :|:|       ||:|:
Db          1 DVPSANANANNQRTAAAKPQANAEAS 26


RESULT 11
S65604
sec c 1 protein - rye (fragment)
C;Species: Secale cereale (rye)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C;Accession: S65604
R;Garcia-Casado, G.; Armentia, A.; Sanchez-Monge, R.; Sanchez, L.M.; Lopez-Otin, C.; Salc
FEBS Lett. 364, 36-40, 1995
A;Title: A major baker's asthma allergen from rye flour is considerably more active than
A;Reference number: S65604; MUID:95269763; PMID:7750539
A;Accession: S65604
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-26 <GAR>

  Query Match             1.8%;  Score 37.5;  DB 2;  Length 26;
  Best Local Similarity   46.7%;  Pred. No. 1.1e+04;
  Matches    7;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        208 GSSVQWSDPLPECRE 222
              | |:  ::|:| |||
Db         11 GKSIS-NNPVPACRE 24


RESULT 12
D45900
complement C3b receptor type 2 - mouse (clone 12) (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: D45900
R;Kurtz, C.B.; O'Toole, E.; Christensen, S.M.; Weis, J.H.
J. Immunol. 144, 3581-3591, 1990
A;Title: The murine complement receptor gene family. IV. Alternative splicing of Cr2 gene
A;Reference number: A45900; MUID:90229754; PMID:2139460
A;Accession: D45900
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tran
A;Molecule type: mRNA
A;Residues: 1-11 <KUR>

  Query Match             1.8%;  Score 37;  DB 2;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 4.3e+03;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        220 CREIYCPAPP 229
              | || |  ||
Db          2 CEEISCDPPP 11
```

RESULT 13
PC1151
equinatoxin 1C - sea anemone (Actinia equina) (fragment)
C;Species: Actinia equina
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PC1151
R;Komatsu, S.; Furukawa, K.; Abe, K.; Hirano, H.; Ueda, M.
Chem. Pharm. Bull. 40, 2873-2875, 1992
A;Title: Isolation and characterization of equinatoxins from the sea anemone Actinia equ
A;Reference number: PC1149; MUID:93099631; PMID:1361161
A;Accession: PC1151
A;Molecule type: protein
A;Residues: 1-20 <KOM>
C;Keywords: toxin

Query Match 1.8%; Score 37; DB 2; Length 20;
Best Local Similarity 37.5%; Pred. No. 8.5e+03;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy        174 IDVPGGILFGATISFS 189
              : | | :: |||::|:
Db          2 VAVAGAVIEGATLTFN 17
```

RESULT 14
B34016
tenebrosin B - sea anemone (Actinia tenebrosa) (fragment)
C;Species: Actinia tenebrosa
C;Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 30-Sep-1993
C;Accession: B34016
R;Norton, R.S.; Bobek, G.; Ivanov, J.O.; Thomson, M.; Fiala-Beer, E.; Moritz, R.L.; Simp
Toxicon 28, 29-41, 1990
A;Title: Purification and characterisation of proteins with cardiac stimulatory and haem
A;Reference number: A34016; MUID:90232538; PMID:1970442
A;Accession: B34016
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <NOR>

Query Match 1.8%; Score 37; DB 2; Length 20;
Best Local Similarity 37.5%; Pred. No. 8.5e+03;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy        174 IDVPGGILFGATISFS 189
              : | | :: |||::|:
Db          2 VAVAGAVIEGATLTFN 17
```

RESULT 15
B64669
hypothetical protein HP1194 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C;Accession: B64669
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: B64669
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-28 <TOM>
A;Cross-references: GB:AE000625; GB:AE000511; NID:g2314349; PIDN:AAD08244.1; PID:g231435

Query Match 1.8%; Score 37; DB 2; Length 28;
Best Local Similarity 36.8%; Pred. No. 1.2e+04;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

```
Qy        130 RPGYRREPSLSPKLTCLQN 148
              || |::    : || |:|:
Db         10 RPPYKKFRLTNDKLFCIQS 28
```

Search completed: November 18, 2003, 07:49:56
Job time : 29 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 18, 2003, 07:42:16 ; Search time 243 Seconds
(without alignments)
286.235 Million cell updates/sec

Title: US-09-623-035-2
Perfect score: 2064
Sequence: 1 MTVARPSVPAALPLLGELPR.....HTCFTLTGLLGTLVTMGLLT 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 149430

Minimum DB seq length: 6
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	64.5	3.1	23	10	US-09-142-043-8	Sequence 8, Appli
2	61	3.0	10	11	US-09-572-404B-3217	Sequence 3217, Ap
3	58	2.8	10	11	US-09-572-404B-3275	Sequence 3275, Ap
4	58	2.8	10	11	US-09-572-404B-3277	Sequence 3277, Ap
5	57	2.8	10	11	US-09-572-404B-3237	Sequence 3237, Ap
6	55	2.7	10	11	US-09-572-404B-3231	Sequence 3231, Ap
7	55	2.7	10	11	US-09-572-404B-3239	Sequence 3239, Ap
8	54	2.6	10	11	US-09-572-404B-3283	Sequence 3283, Ap
9	52	2.5	10	11	US-09-572-404B-3191	Sequence 3191, Ap
10	52	2.5	10	11	US-09-572-404B-3195	Sequence 3195, Ap
11	52	2.5	10	11	US-09-572-404B-3209	Sequence 3209, Ap
12	52	2.5	10	11	US-09-572-404B-3253	Sequence 3253, Ap
13	52	2.5	10	11	US-09-572-404B-3285	Sequence 3285, Ap
14	51	2.5	10	11	US-09-572-404B-3193	Sequence 3193, Ap
15	51	2.5	10	11	US-09-572-404B-3197	Sequence 3197, Ap
16	51	2.5	10	11	US-09-572-404B-3199	Sequence 3199, Ap
17	51	2.5	10	11	US-09-572-404B-3201	Sequence 3201, Ap
18	51	2.5	10	11	US-09-572-404B-3203	Sequence 3203, Ap
19	51	2.5	10	11	US-09-572-404B-3205	Sequence 3205, Ap
20	51	2.5	10	11	US-09-572-404B-3207	Sequence 3207, Ap
21	51	2.5	10	11	US-09-572-404B-3211	Sequence 3211, Ap
22	51	2.5	10	11	US-09-572-404B-3225	Sequence 3225, Ap
23	51	2.5	10	11	US-09-572-404B-3227	Sequence 3227, Ap
24	51	2.5	10	11	US-09-572-404B-3235	Sequence 3235, Ap
25	51	2.5	10	11	US-09-572-404B-3243	Sequence 3243, Ap
26	51	2.5	10	11	US-09-572-404B-3245	Sequence 3245, Ap
27	51	2.5	10	11	US-09-572-404B-3247	Sequence 3247, Ap
28	51	2.5	10	11	US-09-572-404B-3249	Sequence 3249, Ap
29	51	2.5	10	11	US-09-572-404B-3263	Sequence 3263, Ap
30	51	2.5	10	11	US-09-572-404B-3265	Sequence 3265, Ap
31	51	2.5	10	11	US-09-572-404B-3267	Sequence 3267, Ap
32	51	2.5	10	11	US-09-572-404B-3269	Sequence 3269, Ap
33	51	2.5	10	11	US-09-572-404B-3271	Sequence 3271, Ap
34	51	2.5	10	11	US-09-572-404B-3273	Sequence 3273, Ap
35	51	2.5	10	11	US-09-572-404B-3279	Sequence 3279, Ap
36	51	2.5	10	11	US-09-572-404B-3287	Sequence 3287, Ap
37	51	2.5	10	11	US-09-572-404B-3289	Sequence 3289, Ap
38	51	2.5	10	11	US-09-572-404B-3291	Sequence 3291, Ap
39	51	2.5	10	11	US-09-572-404B-3299	Sequence 3299, Ap
40	51	2.5	10	11	US-09-572-404B-3301	Sequence 3301, Ap
41	51	2.5	10	11	US-09-572-404B-3303	Sequence 3303, Ap
42	51	2.5	10	11	US-09-572-404B-3309	Sequence 3309, Ap
43	51	2.5	10	11	US-09-572-404B-3317	Sequence 3317, Ap
44	51	2.5	10	11	US-09-572-404B-3319	Sequence 3319, Ap
45	51	2.5	10	11	US-09-572-404B-3321	Sequence 3321, Ap

ALIGNMENTS

RESULT 1
US-09-142-043-8
; Sequence 8, Application US/09142043
; Patent No. US20020142372A1
; GENERAL INFORMATION:
; APPLICANT: MOSSAKOWSKA, Danuta Ewa Irena
; APPLICANT: EDGE, Colin Michael
; APPLICANT: SMITH, Richard Anthony Godwin
; TITLE OF INVENTION: FRAGMENTS OF CR1 AND THEIR USE
; FILE REFERENCE: 88362/104
; CURRENT APPLICATION NUMBER: US/09/142,043
; CURRENT FILING DATE: 1998-12-01
; EARLIER APPLICATION NUMBER: WO PCT/EP97/00994
; EARLIER FILING DATE: 1997-02-26
; EARLIER APPLICATION NUMBER: GB 96045182.2
; EARLIER FILING DATE: 1996-03-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-142-043-8

Query Match 3.1%; Score 64.5; DB 10; Length 23;
Best Local Similarity 56.5%; Pred. No. 99;
Matches 13; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

```
Qy        257 FTMIGEHSIYCTVNNDE-GEWSG 278
              | ::|| ||| | |:|: | |||
Db          1 FELVGEPSIYSTSNDDQVGIWSG 23
```

RESULT 2
US-09-572-404B-3217
; Sequence 3217, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:

```
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3217
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 278-287 and may interact with
;   OTHER INFORMATION: Sequence 3218 in this patent.
US-09-572-404B-3217

  Query Match             3.0%;  Score 61;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 63;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        278 GPPPECRGKS 287
              ||||||||||
Db          1 GPPPECRGKS 10


RESULT 3
US-09-572-404B-3275
; Sequence 3275, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3275
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 151-160 and may interact with
;   OTHER INFORMATION: Sequence 3276 in this patent.
US-09-572-404B-3275

  Query Match             2.8%;  Score 58;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        151 WSTAVEFCKK 160
              ||||||||||
Db          1 WSTAVEFCKK 10


RESULT 4
US-09-572-404B-3277
; Sequence 3277, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3277
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 361-370 and may interact with
;   OTHER INFORMATION: Sequence 3278 in this patent.
US-09-572-404B-3277

  Query Match             2.8%;  Score 58;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        361 GHTCFTLTGL 370
              ||||||||||
Db          1 GHTCFTLTGL 10


RESULT 5
US-09-572-404B-3237
; Sequence 3237, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3237
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 123-132 and may interact with
;   OTHER INFORMATION: Sequence 3238 in this patent.
US-09-572-404B-3237

  Query Match             2.8%;  Score 57;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        123 TVVEYECRPG 132
              ||||||||||
Db          1 TVVEYECRPG 10


RESULT 6
US-09-572-404B-3231
; Sequence 3231, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3231
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 247-256 and may interact with
;   OTHER INFORMATION: Sequence 3232 in this patent.
US-09-572-404B-3231

  Query Match             2.7%;  Score 55;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        247 QSVTYACNKG 256
              ||||||||||
Db          1 QSVTYACNKG 10


RESULT 7
```

US-09-572-404B-3239
; Sequence 3239, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3239
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in DAF OR CD55 at 358-367 and may interact with
; OTHER INFORMATION: Sequence 3240 in this patent.
US-09-572-404B-3239

Query Match 2.7%; Score 55; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 358 LLSGHTCFTL 367
||||||||||
Db 1 LLSGHTCFTL 10

RESULT 8
US-09-572-404B-3283
; Sequence 3283, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3283
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in DAF OR CD55 at 291-300 and may interact with
; OTHER INFORMATION: Sequence 3284 in this patent.
US-09-572-404B-3283

Query Match 2.6%; Score 54; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 291 KVPPTVQKPT 300
||||||||||
Db 1 KVPPTVQKPT 10

RESULT 9
US-09-572-404B-3191
; Sequence 3191, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3191
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in DAF OR CD55 at 196-205 and may interact with
; OTHER INFORMATION: Sequence 3192 in this patent.
US-09-572-404B-3191

Query Match 2.5%; Score 52; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 196 LFGSTSSFCL 205
||||||||||
Db 1 LFGSTSSFCL 10

RESULT 10
US-09-572-404B-3195
; Sequence 3195, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3195
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in DAF OR CD55 at 316-325 and may interact with
; OTHER INFORMATION: Sequence 3196 in this patent.
US-09-572-404B-3195

Query Match 2.5%; Score 52; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 316 TTTKTTTPNA 325
||||||||||
Db 1 TTTKTTTPNA 10

RESULT 11
US-09-572-404B-3209
; Sequence 3209, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3209
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in DAF OR CD55 at 192-201 and may interact with
; OTHER INFORMATION: Sequence 3210 in this patent.
US-09-572-404B-3209

Query Match 2.5%; Score 52; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 192 TGYKLFGSTS 201
||||||||||

```
Db          1 TGYKLFGSTS 10


RESULT 12
US-09-572-404B-3253
; Sequence 3253, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3253
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 330-339 and may interact with
;   OTHER INFORMATION: Sequence 3254 in this patent.
US-09-572-404B-3253

  Query Match             2.5%;  Score 52;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        330 STPVSRTTKH 339
              ||||||||||
Db          1 STPVSRTTKH 10


RESULT 13
US-09-572-404B-3285
; Sequence 3285, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3285
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 346-355 and may interact with
;   OTHER INFORMATION: Sequence 3286 in this patent.
US-09-572-404B-3285

  Query Match             2.5%;  Score 52;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        346 NKGSGTTSGT 355
              ||||||||||
Db          1 NKGSGTTSGT 10


RESULT 14
US-09-572-404B-3193
; Sequence 3193, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3193
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 348-357 and may interact with
;   OTHER INFORMATION: Sequence 3194 in this patent.
US-09-572-404B-3193

  Query Match             2.5%;  Score 51;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        348 GSGTTSGTTR 357
              ||||||||||
Db          1 GSGTTSGTTR 10


RESULT 15
US-09-572-404B-3197
; Sequence 3197, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3197
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in DAF OR CD55 at 348-357 and may interact with
;   OTHER INFORMATION: Sequence 3198 in this patent.
US-09-572-404B-3197

  Query Match             2.5%;  Score 51;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        348 GSGTTSGTTR 357
              ||||||||||
Db          1 GSGTTSGTTR 10


Search completed: November 18, 2003, 07:48:17
Job time : 243 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 18, 2003, 07:42:11 ; Search time 17 Seconds
(without alignments)
1053.951 Million cell updates/sec

Title: US-09-623-035-2
Perfect score: 2064
Sequence: 1 MTVARPSVPAALPLLGELPR.....HTCFTLTGLLGTLVTMGLLT 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 1987

Minimum DB seq length: 6
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	1.9	27	1	TXA3_ANESU	P01535 anemonia su
2	39.5	1.9	23	1	NEUU_LITCE	P81872 litoria cae
3	38	1.8	27	1	PPS1_DROFU	P01372 drosophila
4	37	1.8	20	1	TENB_ACTTE	P30834 actinia ten
5	37	1.8	25	1	SODC_PAROL	P83129 paralichthy
6	36.5	1.8	27	1	CX2B_CONBE	P81727 conus betul
7	36	1.7	21	1	JAP2_RANJA	P83306 rana japoni
8	36	1.7	23	1	CP23_SPOER	P56683 spodoptera
9	36	1.7	30	1	AATC_RABIT	P12343 oryctolagus
10	36	1.7	30	1	TAT_HV1ZH	P12512 human immun
11	35	1.7	18	1	PPM2_LIMPO	P14216 limulus pol
12	35	1.7	20	1	TENA_ACTTE	P30833 actinia ten
13	34.5	1.7	28	1	SLP1_LEIQH	P80669 leiurus qui
14	34	1.6	23	1	PAP1_SPOEX	P30255 spodoptera
15	34	1.6	23	1	PAP2_SPOEX	P30256 spodoptera
16	34	1.6	23	1	PAP3_SPOEX	P30257 spodoptera
17	34	1.6	28	1	NXL1_BOUAN	P34074 boulengerin
18	33.5	1.6	13	1	BPP1_BOTJA	P01020 bothrops ja
19	33.5	1.6	23	1	SCK1_TITCA	P83243 tityus camb
20	33.5	1.6	30	1	VARG_VIOAR	P58452 viola arven
21	33	1.6	25	1	AU51_LITRA	P82401 litoria ran
22	33	1.6	29	1	VARF_VIOAR	P58451 viola arven
23	33	1.6	30	1	VARB_VIOAR	P58447 viola arven
24	32.5	1.6	26	1	CXO6_CONTU	P58915 conus tulip
25	32.5	1.6	26	1	NIFQ_ENTAG	P28580 enterobacte
26	32.5	1.6	27	1	ACH4_MOUSE	O70174 mus musculu
27	32.5	1.6	28	1	SCX2_BUTSI	P15230 buthus sind
28	32	1.6	12	1	CD11_LITXA	P56245 litoria xan
29	32	1.6	16	1	CXAA_CONPE	P50984 conus penna
30	32	1.6	18	1	PPM1_LIMPO	P14215 limulus pol
31	32	1.6	21	1	GYRA_STRSH	P50071 streptomyce
32	32	1.6	22	1	LANM_STRMU	P80666 streptococc
33	32	1.6	23	1	ALL5_HORSE	P82616 equus cabal
34	32	1.6	27	1	CX7A_CONTE	P24160 conus texti
35	32	1.6	27	1	LSP_APIME	P09355 apis mellif
36	32	1.6	29	1	MULR_ECHML	P81798 echis multi
37	32	1.6	30	1	CXVB_CONER	P58783 conus ermin
38	31.5	1.5	24	1	SCX5_BUTOC	P04097 buthus occi
39	31.5	1.5	29	1	PSAF_SYNP6	P31083 synechococc
40	31.5	1.5	30	1	CYO1_VIOOD	P82230 viola odora
41	31	1.5	24	1	CS33_ARAHY	P80926 arachis hyp
42	31	1.5	25	1	PRLA_ACHLY	P27459 achromobact
43	31	1.5	27	1	IA17_WHEAT	P16852 triticum ae
44	31	1.5	27	1	MDO2_PSEMR	P83447 pseudananas
45	31	1.5	27	1	SODM_DESDE	P11419 desulfovibr

ALIGNMENTS

RESULT 1
TXA3_ANESU
ID TXA3_ANESU STANDARD; PRT; 27 AA.
AC P01535;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurotoxin III (Toxin ATX-III).
OS Anemonia sulcata (Snake-locks sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynantheae; Actiniidae; Anemonia.
OX NCBI_TaxID=6108;
RN [1]
RP SEQUENCE.
RC TISSUE=Cnidoblast;
RX MEDLINE=78084776; PubMed=23311;
RA Martinez G., Kopeyan C., Schweitz H., Lazdunski M.;
RT "Toxin III from Anemonia sulcata: primary structure.";
RL FEBS Lett. 84:247-252(1977).
RN [2]
RP PRELIMINARY SEQUENCE.
RX MEDLINE=78044787; PubMed=21843;
RA Beress L., Wunderer G., Wachter E.;
RT "Amino acid sequence of toxin III from Anemonia sulcata.";
RL Hoppe-Seyler's Z. Physiol. Chem. 358:985-988(1977).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=93343891; PubMed=8102051;
RA Norton R.S., Cross K., Braach-Maksvytis V., Wachter E.;
RT "1H-NMR study of the solution properties and secondary structure of
RT neurotoxin III from the sea anemone Anemonia sulcata.";
RL Biochem. J. 293:545-551(1993).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE=95244415; PubMed=7727358;
RA Manoleras N., Norton R.S.;
RT "Three-dimensional structure in solution of neurotoxin III from the
RT sea anemone Anemonia sulcata.";
RL Biochemistry 33:11051-11061(1994).
CC -!- FUNCTION: BINDS SPECIFICALLY TO THE SODIUM CHANNEL.
CC -!- SUBCELLULAR LOCATION: Secreted; cnidocyst.
CC -!- SIMILARITY: BELONGS TO THE SEA ANEMONE SHORT TOXIN FAMILY.
DR PIR; A91446; TZAZ3.
DR PDB; 1ANS; 31-AUG-94.
KW Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW Cnidocyst; 3D-structure.
FT DISULFID 3 17
FT DISULFID 4 11
FT DISULFID 6 22
FT CONFLICT 22 23 CS -> SC (IN REF. 2).
FT TURN 7 10
SQ SEQUENCE 27 AA; 2938 MW; AA4E261FFAF34A7A CRC64;

Query Match 1.9%; Score 40; DB 1; Length 27;
Best Local Similarity 26.7%; Pred. No. 3.5e+03;
Matches 8; Conservative 2; Mismatches 8; Indels 12; Gaps 1;

```
Qy         28 CLPAVWGDCGLPPDVPNAQPALEGRTSFPE 57
              | |  || |          .    |:  :||
Db          3 CCPCYWGGCPW-----------GQNCYPE 20


RESULT 2
NEUU_LITCE
ID   NEUU_LITCE     STANDARD;      PRT;    23 AA.
AC   P81872;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Neuromedin U-23 (NmU-23).
OS   Litoria caerulea (Green tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=30344;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Skin secretion;
RX   MEDLINE=20138182; PubMed=10671478;
RA   Salmon A.L., Johnsen A.H., Bienert M., McMurray G., Nandha K.A.,
RA   Bloom S.R., Shaw C.;
RT   "Isolation, structural characterization and bioactivity of a novel
RT   neuromedin U analog from the defensive skin secretion of the
RT   Australasian tree frog Litoria caerulea.";
RL   J. Biol. Chem. 275:4549-4554(2000).
CC   -!- FUNCTION: STIMULATE UTERINE SMOOTH MUSCLE CONTRACTION AND CAUSE
CC       SELECTIVE VASOCONSTRICTION.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin.
CC   -!- SIMILARITY: BELONGS TO THE NMU FAMILY.
DR   InterPro; IPR001942; NMU.
DR   Pfam; PF02070; NMU; 1.
DR   PROSITE; PS00967; NMU; 1.
KW   Amidation; Hormone.
FT   MOD_RES       23      23       AMIDATION.
SQ   SEQUENCE   23 AA;  2581 MW;  A94958415CB58DC3 CRC64;

  Query Match             1.9%;  Score 39.5;  DB 1;  Length 23;
  Best Local Similarity   32.0%;  Pred. No. 3.2e+03;
  Matches    8;  Conservative    5;  Mismatches    3;  Indels    9;  Gaps    1;

Qy        173 QIDVPGGILFGATISFSCNTGYKLF 197
              :: ||||::        : || ||
Db          4 EVQVPGGVI---------SNGYFLF 19


RESULT 3
PPS1_DROFU
ID   PPS1_DROFU     STANDARD;      PRT;    27 AA.
AC   P01372;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Paragonial peptide PS-1 (Paragonial peptide C).
GN   PAPC.
OS   Drosophila funebris (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7221;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=77245908; PubMed=1236144;
RA   Baumann H., Wilson K.J., Chen P.S., Humbel R.E.;
RT   "The amino-acid sequence of a peptide (PS-1) from Drosophila
RT   funebris: a paragonial peptide from males which reduces the
RT   receptivity of the female.";
RL   Eur. J. Biochem. 52:521-529(1975).
CC   -!- FUNCTION: REPRESSES FEMALE SEXUAL RECEPTIVITY AND STIMULATES
CC       OVIPOSITION. THIS PEPTIDE HAS A LOW ACTIVITY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLANDS OF MALES
CC       (PARAGONIAL GLAND).
DR   PIR; A01643; PGFF1.
DR   FlyBase; FBgn0004112; Dfun\PapC.
KW   Behavior.
FT   VARIANT       2       2       V -> L (IN 30% OF THE MOLECULES).
SQ   SEQUENCE   27 AA;  2669 MW;  D8083427974AD5D6 CRC64;

  Query Match             1.8%;  Score 38;  DB 1;  Length 27;
  Best Local Similarity   26.9%;  Pred. No. 4.8e+03;
  Matches    7;  Conservative    7;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        303 NVPTTEVSPTSQKTTTKTTTPNAQAT 328
              :||:   :  :|:|       ||:|:
Db          1 DVPSANANANNQRTAAAKPQANAEAS 26


RESULT 4
TENB_ACTTE
ID   TENB_ACTTE     STANDARD;      PRT;    20 AA.
AC   P30834;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tenebrosin B (Fragment).
OS   Actinia tenebrosa (Australian red waratah sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Actinia.
OX   NCBI_TaxID=6105;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=90232538; PubMed=1970442;
RA   Norton R.S., Bobek G., Ivanov J.O., Thomson M., Fiala-Beer E.,
RA   Moritz R.L., Simpson R.J.;
RT   "Purification and characterisation of proteins with cardiac
RT   stimulatory and haemolytic activity from the anemone Actinia
RT   tenebrosa.";
RL   Toxicon 28:29-41(1990).
CC   -!- FUNCTION: This cardiac stimulatory and hemolytic protein is a
CC       channel-forming and/or membrane-penetrating protein.
CC   -!- SUBCELLULAR LOCATION: Secreted; cnidocyst.
CC   -!- SIMILARITY: BELONGS TO THE TENEBROSIN FAMILY.
DR   PIR; B34016; B34016.
KW   Cytolysis; Hemolysis; Toxin; Cnidocyst; Transmembrane.
FT   NON_TER      20      20
SQ   SEQUENCE   20 AA;  1960 MW;  FA32B426009FF5FA CRC64;

  Query Match             1.8%;  Score 37;  DB 1;  Length 20;
  Best Local Similarity   37.5%;  Pred. No. 4.1e+03;
  Matches    6;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        174 IDVPGGILFGATISFS 189
              : | | :: |||::|:
Db          2 VAVAGAVIEGATLTFN 17


RESULT 5
SODC_PAROL
ID   SODC_PAROL     STANDARD;      PRT;    25 AA.
AC   P83129;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) (Fragment).
OS   Paralichthys olivaceus (Flounder).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Pleuronectoidei; Paralichthyidae; Paralichthys.
OX   NCBI_TaxID=8255;
RN   [1]
```

```
RP   SEQUENCE, AND CHARACTERIZATION.
RC   TISSUE=Hepatopancreas;
RX   MEDLINE=21186415; PubMed=11290457;
RA   Osatomi K., Masuda Y., Hara K., Ishihara T.;
RT   "Purification, N-terminal amino acid sequence, and some properties of
RT   Cu, Zn-superoxide dismutase from Japanese flounder (Paralichthys
RT   olivaceus) hepato-pancreas.";
RL   Comp. Biochem. Physiol. 128B:751-760(2001).
CC   -!- FUNCTION: Destroys radicals which are normally produced within the
CC       cells and which are toxic to biological systems (By similarity).
CC   -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC   -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC       similarity).
CC   -!- SUBUNIT: Homotrimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE
CC       FAMILY.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005507; F:copper ion binding activity; ISS.
DR   GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IDA.
DR   GO; GO:0008270; F:zinc ion binding activity; ISS.
DR   GO; GO:0019430; P:removal of superoxide radicals; TAS.
DR   InterPro; IPR001424; SOD_CU_ZN.
DR   Pfam; PF00080; sodcu; 1.
DR   PROSITE; PS00087; SOD_CU_ZN_1; PARTIAL.
DR   PROSITE; PS00332; SOD_CU_ZN_2; PARTIAL.
KW   Antioxidant; Oxidoreductase; Metal-binding; Copper; Zinc.
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2604 MW;  76C4321D740857C5 CRC64;

  Query Match             1.8%;  Score 37;  DB 1;  Length 25;
  Best Local Similarity   77.8%;  Pred. No. 5.2e+03;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        347 KGSGTTSGT 355
              ||:| ||||
Db          9 KGAGETSGT 17


RESULT 6
CX2B_CONBE
ID   CX2B_CONBE     STANDARD;      PRT;    27 AA.
AC   P81727;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Conotoxin BeTXIIb.
OS   Conus betulinus (Beech cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=89764;
RN   [1]
RP   SEQUENCE, MASS SPECTROMETRY, AND HYDROXYLATION.
RC   TISSUE=Venom;
RX   MEDLINE=20058566; PubMed=10591037;
RA   Chen J.-S., Fan C.-X., Hu K.-P., Wei K.-H., Zhong M.-N.;
RT   "Studies on conotoxins of Conus betulinus.";
RL   J. Nat. Toxins 8:341-349(1999).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- MASS SPECTROMETRY: MW=2664.3; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE P-SUPERFAMILY OF CONOTOXINS.
KW   Neurotoxin; Toxin; Hydroxylation.
FT   DISULFID      2     16       BY SIMILARITY.
FT   DISULFID      6     19       BY SIMILARITY.
FT   DISULFID     12     24       BY SIMILARITY.
FT   MOD_RES      13     13       HYDROXYLATION.
SQ   SEQUENCE   27 AA;  2651 MW;  60F0500514C7F1BA CRC64;

  Query Match             1.8%;  Score 36.5;  DB 1;  Length 27;
  Best Local Similarity   44.4%;  Pred. No. 6.1e+03;
  Matches    8;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        178 GGI-LFGATISFSCNTGY 194
              ||:  :| :   |||| |
Db          3 GGVCAYGESCPSSCNTCY 20


RESULT 7
JAP2_RANJA
ID   JAP2_RANJA     STANDARD;      PRT;    21 AA.
AC   P83306;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Japonicin-2.
OS   Rana japonica (Japanese reddish frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8402;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RX   MEDLINE=21826910; PubMed=11835990;
RA   Isaacson T., Soto A., Iwamuro S., Knoop F.C., Conlon J.M.;
RT   "Antimicrobial peptides with atypical structural features from the
RT   skin of the Japanese brown frog Rana japonica.";
RL   Peptides 23:419-425(2002).
CC   -!- FUNCTION: Antibacterial activity against the Gram-negative
CC       bacterium E.coli and the Gram-positive bacterium S.aureus.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin.
CC   -!- MASS SPECTROMETRY: MW=2685.4; METHOD=Electrospray.
KW   Amphibian defense peptide; Antibiotic.
FT   DISULFID     14     21       BY SIMILARITY.
SQ   SEQUENCE   21 AA;  2358 MW;  AB055DF897566BDF CRC64;

  Query Match             1.7%;  Score 36;  DB 1;  Length 21;
  Best Local Similarity   46.7%;  Pred. No. 5e+03;
  Matches    7;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         12 LPLLGELPRLLLLVL 26
              ||:|  ||: | ::|
Db          3 LPMLSILPKALCILL 17


RESULT 8
CP23_SPOER
ID   CP23_SPOER     STANDARD;      PRT;    23 AA.
AC   P56683;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Cardioactive peptide CAP23.
OS   Spodoptera eridania (Southern armyworm).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC   Noctuidae; Amphipyrinae; Spodoptera.
OX   NCBI_TaxID=37547;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=99196260; PubMed=10098624;
RA   Furuya K., Hackett M., Cirelli M.A., Schegg K.M., Wang H.,
RA   Shabanowitz J., Hunt D.F., Schooley D.A.;
RT   "A cardioactive peptide from the southern armyworm, Spodoptera
RT   eridania.";
RL   Peptides 20:53-61(1999).
CC   -!- FUNCTION: HAS EXCITATORY EFFECTS ON A SEMI-ISOLATED HEART FROM
CC       LARVAL MANDUCA SEXTA, CAUSING AN INOTROPIC EFFECT AT LOW
CC       CONCENTRATIONS OF PEPTIDE AND CHRONOTROPIC AND INOTROPIC EFFECTS
CC       AT HIGH DOSES.
CC   -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR   HSSP; O61704; 1B1V.
DR   InterPro; IPR003463; GBP_PSP.
DR   Pfam; PF02425; GBP_PSP; 1.
```

```
DR   ProDom; PD006507; GBP_PSP; 1.
FT   DISULFID      7     19       BY SIMILARITY.
SQ   SEQUENCE   23 AA;  2519 MW;  0A96D72A70855AE0 CRC64;

  Query Match             1.7%;  Score 36;  DB 1;  Length 23;
  Best Local Similarity   47.1%;  Pred. No. 5.5e+03;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    6;  Gaps    1;

Qy        119 FPVGTVVEYECRPGYRR 135
              | ||      | |||:|
Db          3 FAVG------CTPGYQR 13


RESULT 9
AATC_RABIT
ID   AATC_RABIT     STANDARD;      PRT;    30 AA.
AC   P12343;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Aspartate aminotransferase, cytoplasmic (EC 2.6.1.1) (Transaminase A)
DE   (Glutamate oxaloacetate transaminase-1) (Fragment).
GN   GOT1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=85289123; PubMed=4030726;
RA   Kuramitsu S., Inoue K., Kondo K., Aki K., Kagamiyama H.;
RT   "Aspartate aminotransferase isozymes from rabbit liver. Purification
RT   and properties.";
RL   J. Biochem. 97:1337-1345(1985).
CC   -!- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC       L-glutamate.
CC   -!- COFACTOR: Pyridoxal phosphate.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: In eukaryotes there are cytoplasmic, mitochondrial
CC       and chloroplastic isozymes.
CC   -!- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC       AMINOTRANSFERASES.
DR   PIR; A27103; A27103.
DR   HSSP; P00503; 1AJS.
DR   InterPro; IPR004838; NHtransf_1.
DR   PROSITE; PS00105; AA_TRANSFER_CLASS_1; PARTIAL.
KW   Transferase; Aminotransferase; Pyridoxal phosphate.
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3296 MW;  9B76CEEAC2FC5D98 CRC64;

  Query Match             1.7%;  Score 36;  DB 1;  Length 30;
  Best Local Similarity   42.3%;  Pred. No. 7.4e+03;
  Matches   11;  Conservative    2;  Mismatches    7;  Indels    6;  Gaps    2;

Qy         39 PP----DVPNAQPALEGR--TSFPED 58
              ||    :|| ||| |  :    | ||
Db          2 PPSIFAEVPQAQPVLVFKLTADFRED 27


RESULT 10
TAT_HV1ZH
ID   TAT_HV1ZH      STANDARD;      PRT;    30 AA.
AC   P12512;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   TAT protein (Transactivating regulatory protein) (Fragment).
GN   TAT.
OS   Human immunodeficiency virus type 1 (Zaire HZ321 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11692;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89228766; PubMed=2713163;
RA   Srinivasan A., York D., Butler D. Jr., Jannoun-Nasr R., Getchell J.,
RA   McCormick J., Ou C.Y., Myers G., Smith T., Chen E.;
RT   "Molecular characterization of HIV-1 isolated from a serum collected
RT   in 1976: nucleotide sequence comparison to recent isolates and
RT   generation of hybrid HIV.";
RL   AIDS Res. Hum. Retroviruses 5:121-129(1989).
CC   -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC       TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC       ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC       PROMOTER.
CC   -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M15896; AAB53949.1; -.
DR   HIV; M15896; TAT$Z321.
KW   Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW   AIDS.
FT   NON_TER       1      1
SQ   SEQUENCE   30 AA;  3329 MW;  545F848858040A1F CRC64;

  Query Match             1.7%;  Score 36;  DB 1;  Length 30;
  Best Local Similarity   43.8%;  Pred. No. 7.4e+03;
  Matches    7;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        304 VPTTEVSPTSQKTTTK 319
              :|||  :||  | : |
Db          2 LPTTRGNPTGPKESKK 17


RESULT 11
PPM2_LIMPO
ID   PPM2_LIMPO     STANDARD;      PRT;    18 AA.
AC   P14216;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   Polyphemusin II.
OS   Limulus polyphemus (Atlantic horseshoe crab).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC   Limulidae; Limulus.
OX   NCBI_TaxID=6850;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=90110066; PubMed=2514185;
RA   Miyata T., Tokunaga F., Yonega T., Yoshikawa K., Iwanaga S., Niwa M.,
RA   Takao T., Shimonishi Y.;
RT   "Antimicrobial peptides, isolated from horseshoe crab hemocytes,
RT   tachyplesin II, and polyphemusins I and II: chemical structures and
RT   biological activity.";
RL   J. Biochem. 106:663-668(1989).
CC   -!- FUNCTION: SIGNIFICANTLY INHIBITS THE GROWTH OF GRAM-NEGATIVE AND
CC       GRAM-POSITIVE BACTERIA.
CC   -!- TISSUE SPECIFICITY: HEMOCYTES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYPLESIN/POLYPHEMUSIN FAMILY.
DR   PIR; JU0125; JU0125.
KW   Antibiotic; Amidation.
FT   DISULFID      4     17       BY SIMILARITY.
FT   DISULFID      8     13       BY SIMILARITY.
FT   MOD_RES      18     18       AMIDATION.
SQ   SEQUENCE   18 AA;  2431 MW;  E402A109D2923504 CRC64;

  Query Match             1.7%;  Score 35;  DB 1;  Length 18;
```

```
  Best Local Similarity   36.8%;  Pred. No. 5e+03;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    8;  Gaps    1;

Qy         80 ICLKGSQWSDIEEFCNRSC 98
              :| ||         || | |
Db          7 VCYKG--------FCYRKC 17


RESULT 12
TENA_ACTTE
ID   TENA_ACTTE     STANDARD;      PRT;     20 AA.
AC   P30833;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tenebrosin A (Fragment).
OS   Actinia tenebrosa (Australian red waratah sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Actinia.
OX   NCBI_TaxID=6105;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=90232538; PubMed=1970442;
RA   Norton R.S., Bobek G., Ivanov J.O., Thomson M., Fiala-Beer E.,
RA   Moritz R.L., Simpson R.J.;
RT   "Purification and characterisation of proteins with cardiac
RT   stimulatory and haemolytic activity from the anemone Actinia
RT   tenebrosa.";
RL   Toxicon 28:29-41(1990).
CC   -!- FUNCTION: This cardiac stimulatory and hemolytic protein is a
CC       channel-forming and/or membrane-penetrating protein.
CC   -!- SUBCELLULAR LOCATION: Secreted; cnidocyst.
CC   -!- SIMILARITY: BELONGS TO THE TENEBROSIN FAMILY.
DR   PIR; A34016; A34016.
KW   Cytolysis; Hemolysis; Toxin; Cnidocyst; Transmembrane.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  1974 MW;  FA32AC8BDAAFF5FA CRC64;

  Query Match             1.7%;  Score 35;  DB 1;  Length 20;
  Best Local Similarity   46.2%;  Pred. No. 5.6e+03;
  Matches    6;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        176 VPGGILFGATISF 188
              | | :: |||:|
Db          4 VAGAVIEGATLTF 16


RESULT 13
SLP1_LEIQH
ID   SLP1_LEIQH     STANDARD;      PRT;     28 AA.
AC   P80669;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leiuropeptide I.
OS   Leiurus quinquestriatus hebraeus (Yellow scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6884;
RN   [1]
RP   SEQUENCE, AND STRUCTURE BY NMR.
RC   TISSUE=Venom;
RX   MEDLINE=97411504; PubMed=9266482;
RA   Buisine E., Wieruszeski J.-M., Lippens G., Wouters D., Tartar A.,
RA   Sautiere P.;
RT   "Characterization of a new family of toxin-like peptides from the
RT   venom of the scorpion Leiurus quinquestriatus hebraeus. 1H-NMR
RT   structure of leiuropeptide II.";
RL   J. Pept. Res. 49:545-555(1997).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE SCORPION LEIUROTOXIN FAMILY.
DR   HSSP; Q9NJP7; 1DU9.
KW   Toxin.
FT   DISULFID      3     19
FT   DISULFID      6     24
FT   DISULFID     10     26
SQ   SEQUENCE   28 AA;  2954 MW;  5F72AD78BD39BE1B CRC64;

  Query Match             1.7%;  Score 34.5;  DB 1;  Length 28;
  Best Local Similarity   39.1%;  Pred. No. 8.7e+03;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    9;  Gaps    2;

Qy        220 CREIYCP-------APPQIDNGI 235
              | |  ||       | |  |||:
Db          3 CEE--CPMHCKGKNAKPTCDNGV 23


RESULT 14
PAP1_SPOEX
ID   PAP1_SPOEX     STANDARD;      PRT;     23 AA.
AC   P30255;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Paralytic peptide I (PP I).
OS   Spodoptera exigua (Beet armyworm).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC   Noctuidae; Amphipyrinae; Spodoptera.
OX   NCBI_TaxID=7107;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Hemolymph;
RX   MEDLINE=91302298; PubMed=2071576;
RA   Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA   Quistad G.B.;
RT   "Isolation and identification of paralytic peptides from hemolymph of
RT   the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT   Heliothis virescens.";
RL   J. Biol. Chem. 266:12873-12877(1991).
CC   -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC       LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC       HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC   -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR   PIR; C39855; C39855.
DR   HSSP; O61704; 1B1V.
DR   InterPro; IPR003463; GBP_PSP.
DR   Pfam; PF02425; GBP_PSP; 1.
DR   ProDom; PD006507; GBP_PSP; 1.
KW   Hemolymph.
FT   DISULFID      7     19       BY SIMILARITY.
SQ   SEQUENCE   23 AA;  2451 MW;  0A96D1F600855AE0 CRC64;

  Query Match             1.6%;  Score 34;  DB 1;  Length 23;
  Best Local Similarity   71.4%;  Pred. No. 7.6e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        129 CRPGYRR 135
              | |||:|
Db          7 CTPGYQR 13


RESULT 15
PAP2_SPOEX
ID   PAP2_SPOEX     STANDARD;      PRT;     23 AA.
AC   P30256;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Paralytic peptide II (PP II).
OS   Spodoptera exigua (Beet armyworm).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC   Noctuidae; Amphipyrinae; Spodoptera.
OX   NCBI_TaxID=7107;
```

```
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Hemolymph;
RX   MEDLINE=91302298; PubMed=2071576;
RA   Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA   Quistad G.B.;
RT   "Isolation and identification of paralytic peptides from hemolymph of
RT   the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT   Heliothis virescens.";
RL   J. Biol. Chem. 266:12873-12877(1991).
CC   -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC       LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC       HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC   -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR   PIR; D39855; D39855.
DR   HSSP; O61704; 1B1V.
DR   InterPro; IPR003463; GBP_PSP.
DR   Pfam; PF02425; GBP_PSP; 1.
DR   ProDom; PD006507; GBP_PSP; 1.
KW   Hemolymph.
FT   DISULFID      7     19       BY SIMILARITY.
SQ   SEQUENCE   23 AA;  2477 MW;  0A96CB4600855AE0 CRC64;

  Query Match             1.6%;  Score 34;  DB 1;  Length 23;
  Best Local Similarity   71.4%;  Pred. No. 7.6e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        129 CRPGYRR 135
              | |||:|
Db          7 CTPGYQR 13


Search completed: November 18, 2003, 07:42:42
Job time : 19 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 18, 2003, 07:42:16 ; Search time 56 Seconds
(without alignments)
1755.679 Million cell updates/sec

Title: US-09-623-035-2
Perfect score: 2064
Sequence: 1 MTVARPSVPAALPLLGELPR.....HTCFTLTGLLGTLVTMGLLT 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 16438

Minimum DB seq length: 6
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	2.3	30	5	Q27545	Q27545 crithidia f
2	45.5	2.2	25	6	Q9BGI9	Q9bgi9 cheirogaleu
3	44.5	2.2	22	11	Q9R1C1	Q9r1c1 mus musculu
4	43	2.1	29	15	O71992	O71992 human endog
5	40.5	2.0	20	4	Q96T45	Q96t45 homo sapien
6	40.5	2.0	26	11	Q9JMC5	Q9jmc5 rattus norv
7	40.5	2.0	30	11	Q9QV95	Q9qv95 cavia (guin
8	40	1.9	20	3	Q9URC1	Q9urc1 phanerochae
9	40	1.9	27	13	O57555	O57555 lampetra pl
10	40	1.9	27	13	O57554	O57554 lampetra pl
11	40	1.9	30	11	Q9JK06	Q9jk06 mus musculu
12	39.5	1.9	26	5	Q8N066	Q8n066 plasmodium
13	39	1.9	25	6	Q9BGJ0	Q9bgj0 eulemur mac
14	39	1.9	27	12	Q9IJ66	Q9ij66 hepatitis c
15	39	1.9	28	6	Q9XS67	Q9xs67 bos taurus
16	39	1.9	30	13	Q9PRP7	Q9prp7 echis carin
17	38.5	1.9	30	13	Q9PRP8	Q9prp8 echis carin
18	38	1.8	19	6	Q8MJ41	Q8mj41 bos taurus
19	38	1.8	23	11	Q64014	Q64014 mus sp. inh
20	38	1.8	23	11	Q9R231	Q9r231 rattus norv
21	38	1.8	24	16	Q9K8M1	Q9k8m1 bacillus ha
22	38	1.8	25	12	O11469	O11469 hepatitis c
23	38	1.8	26	6	Q9TRM6	Q9trm6 bos taurus
24	38	1.8	27	5	Q23745	Q23745 ctenodrilus
25	38	1.8	30	4	Q96D69	Q96d69 homo sapien
26	37.5	1.8	25	2	O07916	O07916 mycobacteri
27	37.5	1.8	26	10	Q9S8H2	Q9s8h2 secale cere
28	37.5	1.8	27	12	Q9QIB5	Q9qib5 hepatitis c
29	37.5	1.8	27	12	Q9QIC1	Q9qic1 hepatitis c
30	37.5	1.8	27	12	Q9QIB3	Q9qib3 hepatitis c
31	37.5	1.8	27	12	Q9QIB9	Q9qib9 hepatitis c
32	37.5	1.8	27	12	Q9QIB2	Q9qib2 hepatitis c
33	37.5	1.8	27	12	Q9QIC7	Q9qic7 hepatitis c
34	37.5	1.8	27	12	Q9QIB6	Q9qib6 hepatitis c
35	37.5	1.8	27	12	Q9QIC6	Q9qic6 hepatitis c
36	37.5	1.8	27	12	Q9QIB0	Q9qib0 hepatitis c
37	37.5	1.8	27	12	Q9QIB4	Q9qib4 hepatitis c
38	37.5	1.8	27	12	Q9QIC4	Q9qic4 hepatitis c
39	37.5	1.8	27	12	Q9QIB7	Q9qib7 hepatitis c
40	37.5	1.8	27	12	Q9QIC0	Q9qic0 hepatitis c
41	37.5	1.8	27	12	Q9QIC9	Q9qic9 hepatitis c
42	37.5	1.8	27	12	Q9QIB8	Q9qib8 hepatitis c
43	37.5	1.8	27	12	Q9QIC5	Q9qic5 hepatitis c
44	37.5	1.8	30	12	Q9QS21	Q9qs21 hepatitis b
45	37.5	1.8	30	12	Q9QS15	Q9qs15 hepatitis b

ALIGNMENTS

RESULT 1
Q27545
ID Q27545 PRELIMINARY; PRT; 30 AA.
AC Q27545;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Reiske iron-sulfur protein precursor (EC 1.10.2.2) (Fragment).
OS Crithidia fasciculata.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.
OX NCBI_TaxID=5656;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96096683; PubMed=7495871;
RA Priest J.W., Hajduk S.L.;
RT "The trypanosomatid Rieske iron-sulfur proteins have a cleaved
RT presequence that may direct mitochondrial import.";
RL Biochim. Biophys. Acta 1269:201-204(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93015890; PubMed=1328195;
RA Priest J.W., Hajduk S.L.;
RT "Cytochrome c reductase purified from Crithidia fasciculata contains
RT an atypical cytochrome c1.";
RL J. Biol. Chem. 267:20188-20195(1992).
DR EMBL; U28865; AAA89057.1; -.
KW Mitochondrion; Oxidoreductase; Transit peptide.
FT TRANSIT 1 17 MITOCHONDRION.
FT CHAIN 18 >30 REISKE IRON-SULFUR PROTEIN.
FT NON_TER 30 30
SQ SEQUENCE 30 AA; 3399 MW; DE41365941814501 CRC64;

Query Match 2.3%; Score 48; DB 5; Length 30;
Best Local Similarity 44.4%; Pred. No. 2.2e+03;
Matches 12; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

```
Qy        319 KTTTPNAQATRSTPVSRTTKHFHETTP 345
              :| |   ||||:  ||  |   |||
Db          4 RTFTTAFQATRAARVSLLVKQLEGTTP 30
```

```
RESULT 2
Q9BGI9
ID   Q9BGI9      PRELIMINARY;      PRT;    25 AA.
AC   Q9BGI9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ATP synthase beta subunit (Fragment).
OS   Cheirogaleus medius (Fat-tailed dwarf lemur).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC   Cheirogaleus.
OX   NCBI_TaxID=9460;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21100352; PubMed=11156996;
RA   Schmitz J., Ohme M., Zischler H.;
RT   "SINE Insertions in Cladistic Analyses and the Phylogenetic
RT   Affiliations of Tarsius bancanus to Other Primates.";
RL   Genetics 157:777-784(2001).
DR   EMBL; AF278741; AAK13316.1; -.
FT   NON_TER       1      1
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2924 MW;  4DE23CB57E8C2A4B CRC64;

  Query Match             2.2%;  Score 45.5;  DB 6;  Length 25;
  Best Local Similarity   47.8%;  Pred. No. 2.9e+03;
  Matches   11;  Conservative    4;  Mismatches    5;  Indels    3;  Gaps    1;

Qy         14 LLGEL---PRLLLLVLLCLPAVW 33
              |:||:   |  ||| :| || :|
Db          3 LMGEVPSKPNNLLLFMLRLPNLW 25


RESULT 3
Q9R1C1
ID   Q9R1C1      PRELIMINARY;      PRT;    22 AA.
AC   Q9R1C1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Decay accelerating factor (Fragment).
GN   DAF.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=99348049; PubMed=10417349;
RA   Harris C.L., Rushmere N.K., Morgan B.P.;
RT   "Molecular and functional analysis of mouse decay accelerating
RT   factor.";
RL   Biochem. J. 341:821-829(1999).
DR   EMBL; AF143541; AAD51449.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   22 AA;  2371 MW;  3DAD8BF3A02D33DB CRC64;

  Query Match             2.2%;  Score 44.5;  DB 11;  Length 22;
  Best Local Similarity   54.5%;  Pred. No. 3e+03;
  Matches   12;  Conservative    1;  Mismatches    8;  Indels    1;  Gaps    1;

Qy        361 GHTC-FTLTGLLGTLVTMGLLT 381
              ||||  ||| |   |  :| ||
Db          1 GHTCLITLTVLHAMLSLIGYLT 22


RESULT 4
O71992
ID   O71992      PRELIMINARY;      PRT;    29 AA.
AC   O71992;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Reverse transcriptase (Fragment).
OS   Human endogenous retrovirus.
OC   Viruses; Retroid viruses; Retroviridae.
OX   NCBI_TaxID=11827;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rose T., Schultz E.R., Henikoff J.G., Pietrokovski S., McCallum C.M.,
RA   Henikoff S.;
RT   "Consensus-degenerate hybrid oligonucleotide primers for amplification
RT   of distantly-related sequences.";
RL   Nucleic Acids Res. 0:0-0(1998).
DR   EMBL; AF050505; AAC05560.1; -.
FT   NON_TER       1      1
FT   NON_TER      29     29
SQ   SEQUENCE   29 AA;  3303 MW;  CD65212FE4DB04AB CRC64;

  Query Match             2.1%;  Score 43;  DB 15;  Length 29;
  Best Local Similarity   40.9%;  Pred. No. 5.8e+03;
  Matches    9;  Conservative    3;  Mismatches    8;  Indels    2;  Gaps    1;

Qy        109 LKQPYITQNYFPVGTVVEYECR 130
              |  | | | |  ||  :|: |:
Db          1 LNSPTICQTY--VGQAIEFTCK 20


RESULT 5
Q96T45
ID   Q96T45      PRELIMINARY;      PRT;    20 AA.
AC   Q96T45;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MER receptor tyrosine kinase (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20517330; PubMed=11062461;
RA   Gal A., Li Y., Thompson D.A., Weir J., Orth U., Jacobson S.G.,
RA   Apfelstedt-Sylla E., Vollrath D.;
RT   "Mutations in MERTK, the human orthologue of the RCS rat retinal
RT   dystrophy gene, cause retinitis pigmentosa.";
RL   Nat. Genet. 26:270-271(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gal A., Li Y., Thompson D.A., Weir J., Orth U., Jacobson S.G.,
RA   Apfelstedt-Sylla E., Vollrath D.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF366903; AAK54121.1; -.
KW   Kinase; Receptor.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2232 MW;  A853BEF7EECE2910 CRC64;

  Query Match             2.0%;  Score 40.5;  DB 4;  Length 20;
  Best Local Similarity   44.0%;  Pred. No. 6e+03;
  Matches   11;  Conservative    2;  Mismatches    3;  Indels    9;  Gaps    1;

Qy          9 PAALPLLGELPRLLLLVLLCLPAVW 33
              || ||          ||: | |||:|
Db          3 PAPLP---------LLLGLFLPALW 18


RESULT 6
Q9JMC5
ID   Q9JMC5      PRELIMINARY;      PRT;    26 AA.
AC   Q9JMC5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
```

```
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   Decay-accelarating factor (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20130146; PubMed=10663575;
RA   Miwa T., Okada N., Okada H.;
RT   "Alternative exon usage in the 3'region of a single gene generates
RT   glycosylphophatidylinositol-anchored and transmembrane forms of rat
RT   decay-accelerating factor.";
RL   Immunogenetics 51:129-137(2000).
DR   EMBL; AB032396; BAA92772.1; -.
FT   NON_TER       1      1
FT   NON_TER      26     26
SQ   SEQUENCE   26 AA;  2906 MW;  9EEDAEB6206401B3 CRC64;

  Query Match             2.0%;  Score 40.5;  DB 11;  Length 26;
  Best Local Similarity   34.6%;  Pred. No. 8.4e+03;
  Matches    9;  Conservative    7;  Mismatches    9;  Indels    1;  Gaps    1;

Qy        328 TRSTPVSRTT-KHFHETTPNKGSGTT 352
              |:  |||:|| :|   |: ::|   :
Db          1 TQHVPVSKTTVRHPTRTSKDRGESNS 26


RESULT 7
Q9QV95
ID   Q9QV95      PRELIMINARY;      PRT;    30 AA.
AC   Q9QV95;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Organ of CORTI protein (Fragment).
OS   Cavia (guinea pigs).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae.
OX   NCBI_TaxID=10140;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93163013; PubMed=8432690;
RA   Thalmann I., Suzuki H., McCourt D.W., Comegys T.H., Thalmann R.;
RT   "Partial amino acid sequences of organ of Corti proteins OCP1 and
RT   OCP2: a progress report.";
RL   Hear. Res. 64:191-198(1993).
DR   InterPro; IPR001810; F-box.
DR   PROSITE; PS50181; FBOX; 1.
SQ   SEQUENCE   30 AA;  3241 MW;  522DC63FFF68A8FC CRC64;

  Query Match             2.0%;  Score 40.5;  DB 11;  Length 30;
  Best Local Similarity   66.7%;  Pred. No. 1e+04;
  Matches   12;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy         15 LGELPRLLLLVLLC-LPA 31
              | |||  ||| ||  |||
Db          4 LAELPEALLLRLLAQLPA 21


RESULT 8
Q9URC1
ID   Q9URC1      PRELIMINARY;      PRT;    20 AA.
AC   Q9URC1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1,4-beta-D-glucan glucanohydrolase (EC 3.2.1.4) (Fragment).
OS   Phanerochaete chrysosporium.
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC   Aphyllophorales; Corticiaceae; Phanerochaete.
OX   NCBI_TaxID=5306;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92068458; PubMed=1367687;
RA   Uzcategui E., Johansson G., Ek B., Pettersson G.;
RL   J. Biotechnol. 21:143-159(1991).
DR   InterPro; IPR000254; CBD_fungal.
DR   Pfam; PF00734; CBM_1; 1.
SQ   SEQUENCE   20 AA;  1939 MW;  3607F893BD532367 CRC64;

  Query Match             1.9%;  Score 40;  DB 3;  Length 20;
  Best Local Similarity   85.7%;  Pred. No. 6.6e+03;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         31 AVWGDCG 37
              |||| ||
Db          2 AVWGQCG 8


RESULT 9
O57555
ID   O57555      PRELIMINARY;      PRT;    27 AA.
AC   O57555;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Homeobox protein LpHox9C (Fragment).
OS   Lampetra planeri (Brook lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Lampetra.
OX   NCBI_TaxID=7750;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98358009; PubMed=9694633;
RA   Sharman A.C., Holland P.W.;
RT   "Estimation of Hox gene cluster number in lampreys.";
RL   Int. J. Dev. Biol. 42:617-620(1998).
DR   EMBL; AF044811; AAC03015.1; -.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2869 MW;  9F2AE2989CA319E1 CRC64;

  Query Match             1.9%;  Score 40;  DB 13;  Length 27;
  Best Local Similarity   72.7%;  Pred. No. 9.7e+03;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         42 VPNAQPALEGR 52
              ||:| ||| ||
Db          5 VPDAGPALRGR 15


RESULT 10
O57554
ID   O57554      PRELIMINARY;      PRT;    27 AA.
AC   O57554;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Homeobox protein LpHox9B (Fragment).
OS   Lampetra planeri (Brook lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Lampetra.
OX   NCBI_TaxID=7750;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98358009; PubMed=9694633;
RA   Sharman A.C., Holland P.W.;
RT   "Estimation of Hox gene cluster number in lampreys.";
RL   Int. J. Dev. Biol. 42:617-620(1998).
DR   EMBL; AF044810; AAC03014.1; -.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2896 MW;  7517E29A4CA319FB CRC64;
```

```
  Query Match             1.9%;  Score 40;  DB 13;  Length 27;
  Best Local Similarity   72.7%;  Pred. No. 9.7e+03;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         42 VPNAQPALEGR 52
              ||:| ||| ||
Db          5 VPDAGPALRGR 15


RESULT 11
Q9JK06
ID   Q9JK06      PRELIMINARY;      PRT;    30 AA.
AC   Q9JK06;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Fructose-1,6-bisphosphatase (EC 3.1.3.11) (Fragment).
GN   FBP2 OR FBPASE 2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=20237676; PubMed=10773464;
RA   Tillmann H., Stein S., Liehr T., Eschrich K.;
RT   "Structure and chromosomal localization of the human and mouse muscle
RT   fructose-1,6-bisphosphatase genes.";
RL   Gene 247:241-253(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skeletal muscle;
RA   Stein S.;
RT   "Mouse liver fructose-1,6-Bisphosphatase: Gene structure,
RT   transcriptional start point, chromosomal localization, cDNA cloning,
RT   characterization of the recombinant protein, and analysis of tissue-
RT   specific expression.";
RL   Arch. Biochem. Biophys. 0:0-0(0).
DR   EMBL; AJ243023; CAB90670.1; -.
DR   HSSP; P00637; 1BK4.
DR   MGD; MGI:95491; Fbp2.
DR   InterPro; IPR000146; In_FB_phphtase.
DR   Pfam; PF00316; FBPase; 1.
DR   ProDom; PD001491; In_FB_phphtase; 1.
KW   Hydrolase.
FT   NON_TER       1      1
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3014 MW;  55146C35C817E0C1 CRC64;

  Query Match             1.9%;  Score 40;  DB 11;  Length 30;
  Best Local Similarity   42.1%;  Pred. No. 1.1e+04;
  Matches    8;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        193 GYKLFGSTSSFCLISGSSV 211
              || |:|| :   | :|  |
Db          5 GYALYGSATLVALSTGQGV 23


RESULT 12
Q8N066
ID   Q8N066      PRELIMINARY;      PRT;    26 AA.
AC   Q8N066;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   BAEBL (Fragment).
GN   BAEBL.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21996440; PubMed=12000842;
RA   Blair P.L., Witney A., Haynes J.D., Moch J.K., Carucci D.J.,
RA   Adams J.H.;
RT   "Transcripts of developmentally regulated Plasmodium falciparum genes
RT   quantified by real-time RT-PCR.";
RL   Nucleic Acids Res. 30:2224-2231(2002).
DR   EMBL; AF461093; AAM33517.1; -.
FT   NON_TER       1      1
FT   NON_TER      26     26
SQ   SEQUENCE   26 AA;  3020 MW;  D8C1933CAEE37D17 CRC64;

  Query Match             1.9%;  Score 39.5;  DB 5;  Length 26;
  Best Local Similarity   29.6%;  Pred. No. 1e+04;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    5;  Gaps    1;

Qy        253 CNKGFTMIGEHSIYCTVNNDEGEWSGP 279
              ||  ::|      ||| :::     ||
Db          4 CNNEYSM-----EYCTYSDERNSSPGP 25


RESULT 13
Q9BGJ0
ID   Q9BGJ0      PRELIMINARY;      PRT;    25 AA.
AC   Q9BGJ0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ATP synthase beta subunit (Fragment).
OS   Eulemur macaco macaco (Black lemur).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Eulemur.
OX   NCBI_TaxID=30603;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21100352; PubMed=11156996;
RA   Schmitz J., Ohme M., Zischler H.;
RT   "SINE Insertions in Cladistic Analyses and the Phylogenetic
RT   Affiliations of Tarsius bancanus to Other Primates.";
RL   Genetics 157:777-784(2001).
DR   EMBL; AF278740; AAK13315.1; -.
FT   NON_TER       1      1
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  3023 MW;  4DE23CB5658D9A4B CRC64;

  Query Match             1.9%;  Score 39;  DB 6;  Length 25;
  Best Local Similarity   53.3%;  Pred. No. 1.1e+04;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         19 PRLLLLVLLCLPAVW 33
              |  ||| :| || :|
Db         11 PNNLLLFMLRLPNLW 25


RESULT 14
Q9IJ66
ID   Q9IJ66      PRELIMINARY;      PRT;    27 AA.
AC   Q9IJ66;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Envelope protein (Genome polyprotein) (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Alberto S.-F.;
RT   "Influence of the dynamics of Hepatitis C virus quasiespecies in the
RT   histological outcome of liver transplantation.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF221229; AAF77791.1; -.
DR   InterPro; IPR002531; HCV_NS1.
```

```
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2776 MW;  CA21BB0A7FB6C40B CRC64;

  Query Match             1.9%;  Score 39;  DB 12;  Length 27;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+04;
  Matches   10;  Conservative    0;  Mismatches    6;  Indels    4;  Gaps    1;

Qy        352 TSGTTRLLSGHTCFTLTGLL 371
              | ||     ||  | || ||
Db          4 TGGTV----GHNTFRLTSLL 19


RESULT 15
Q9XS67
ID   Q9XS67      PRELIMINARY;      PRT;    28 AA.
AC   Q9XS67;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Very low density lipoprotein receptor (Fragment).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99018030; PubMed=9799803;
RA   Magrane J., Reina M., Pagan R., Luna A., Casaroli-Marano R.P.,
RA   Angelin B., Gafvels M., Vilaro S.;
RT   "Bovine aortic endothelial cells express a variant of the very low
RT   density lipoprotein receptor that lacks the O-linked sugar domain.";
RL   J. Lipid Res. 39:2172-2181(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99283875; PubMed=10356983;
RA   Magrane J., Casaroli-Marano R.P., Reina M., Gafvels M., Vilaro S.;
RT   "The role of O-linked sugars in determining the very low density
RT   lipoprotein receptor stability or release from the cell.";
RL   FEBS Lett. 451:56-62(1999).
DR   EMBL; AF034420; AAD31591.1; -.
KW   Lipoprotein; Receptor.
FT   NON_TER       1      1
FT   NON_TER      28     28
SQ   SEQUENCE   28 AA;  2855 MW;  A407207D44718A54 CRC64;

  Query Match             1.9%;  Score 39;  DB 6;  Length 28;
  Best Local Similarity   47.4%;  Pred. No. 1.3e+04;
  Matches    9;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        295 TVQKPTTVNVPTTEVSPTS 313
              ||    | :  | |:||||
Db          4 TVNYNETKDTNTAEISPTS 22


Search completed: November 18, 2003, 07:49:22
Job time : 58 secs
```